REMARKS/ARGUMENTS

Status of the Claims

Claims 1-3, 6-12, and 38-47 are pending in the current application. Claims 4, 5, and 13-37 have been cancelled without prejudice or disclaimer. Claims 13-37 are cancelled due to the restriction requirement. Claims 1-3, 6, 11, and 12 have been amended. Support for the amendment to claim 1 can be found in original claim 5. New claims 38-47 have been added. Support for the new claims is found in original claim 5. Support for new claims 46 and 47 is further found on lines 27-28 of page 7 of the specification. The specification has been amended to correct obvious errors. No new matter has been added by way of amendment. Reexamination and reconsideration are respectfully requested.

The Objection to the Specification Should be Withdrawn

The specification has been objected to on the grounds that a word is misspelled on line 21 of page 35. This paragraph has been amended to correct this error, thereby obviating the objection.

The Rejection Under 35 U.S.C § 112, Second Paragraph, Should be Withdrawn

Claims 1-12 have been rejected under 35 U.S.C. § 112, second paragraph, on the grounds that they are indefinite. Claims 4 and 5 have been cancelled to expedite prosecution, thereby rendering the rejection of this claim moot. It is respectfully submitted that the rejection should not be applied to claims 1-3 and 6-12 as amended, for the reasons described below.

Claim 1 has been rejected on the grounds that the term "genetically modified" is indefinite. A description of genetic modifications of plants that are encompassed by the invention are described on lines 15 of page 7 through line 30 of page 8, and line 17 of page 22 through line 11 of page 23. Accordingly, one of skill in the art, when reading the claims in light of the supporting specification, would be able to ascertain the meaning and scope of the claim.

Claim 1 has further been rejected on the grounds that the term "protein storage tissue" is indefinite. Plant protein storage tissues are described on line 14 of page 1 and lines 12-13 of

page 2 of the specification. Accordingly, the metes and bounds of the claim as written would be clear to one of skill in the art.

Claim 2 has been rejected on the grounds that it is indefinite for reciting that the protein storage tissue is seed. Claim 2 has been amended to recite that the protein storage tissue is seed tissue, thereby obviating the rejection.

Similarly, claim 3 has been rejected on the grounds that it is indefinite for reciting that the protein storage tissue is selected from tubers, roots and leaves. Claim 3 has been amended to recite that the protein storage tissue is selected from tuber tissue, root tissue or leave tissue, thereby obviating the rejection.

Claim 4 has been rejected on the grounds that the term "vacuolar processing enzymes" is indefinite. Claim 4 has been cancelled to expedite prosecution, thereby rendering the rejection of this claim moot.

Claim 6 has been rejected on the grounds that the term "sequence variant" is indefinite. Variants of the proteins of the invention are described on line 24 of page 16 to line 24 of page 17 of the specification, and the structural and functional limitations of these variants are recited in the claims. Accordingly, the metes and bounds of this term would be clear to one of skill in the art. Nevertheless, claim 6 has been amended to delete this term, thereby rendering the rejection on this grounds moot.

Claim 6 has further been rejected on the grounds that the term "protease activity" is indefinite. This term is defined on lines 13-17 of page 5, and assays for protein activity are described in the specification and the references cited therein, for example, lines 5-10 of page 3 and lines 12-13 of page 10. Accordingly, the metes and bounds of the claim as written would be clear to one of skill in the art. Nevertheless, claim 6 has been amended to delete this term, thereby rendering the rejection on this grounds moot.

Claim 6 has also been rejected on the grounds that the term "stringent conditions" is indefinite. A description of stringent conditions to be used for hybridization may be found on lines 6-29 of page 20 of the specification. Accordingly, the metes and bounds of the claim as written would be clear to one of skill in the art. Nevertheless, claim 6 has been amended to delete this term, thereby rendering the rejection on this grounds moot.

Claim 11 has been rejected on the grounds that the term "transformed" is indefinite because it is not clear that the seed encompasses the transgene that was originally introduced into the parent plant. Claim 11 has been amended to recite that the transformed seed is genetically modified to reduce or eliminate the activity of one or more proteases selected from the group consisting of α-vacuolar processing enzyme, β-vacuolar processing enzyme, γ-vacuolar processing enzyme, ε-vacuolar processing enzyme, aspartic protease AP1, and aspartic protease AP2, thereby obviating the rejection.

Claim 12 has been rejected on the grounds that a promoter cannot be operably linked to a polypeptide. Claim 12 has been amended to recite that the promoter is operably linked to a nucleotide sequence encoding a polypeptide of interest, thereby obviating the rejection.

In view of the above arguments and amendments, all grounds for rejection under 35 U.S.C. § 112, second paragraph, have been obviated overcome. Reconsideration and withdrawal of the rejection are therefore respectfully requested.

The Rejection Under 35 U.S.C. § 112, First Paragraph, Should be Withdrawn

Claims 1-12 have been rejected under 35 U.S.C. § 112, first paragraph, on the grounds that they lack sufficient written description. Claims 4 and 5 have been cancelled, rendering the rejection of these claims moot. The rejection is respectfully traversed as applied to claims 1-3 and 6-12 for the reasons described below. It is also respectfully submitted that the rejection should not be applied to new claims 38-47.

The Examiner argues that the claims lack sufficient written description because they recite an expression cassette encoding a polypeptide of interest but do not provide the structural features common to the polypeptides of interest to be expressed. Applicants note that a large number of examples of polypeptides of interest are provided in the specification on line 28 of page 11 through line 24 of page 12.

Furthermore, the point of novelty of the present invention is not the expression of a genus of novel polypeptides of interest. Rather, the point of novelty of the present invention is the reduction or elimination of the activity of proteases in the protein storage tissue in a plant. The

Guidelines for Examination of Patent Applications Under the 35 U.S.C. § 112, ¶ 1, "Written Description Requirement," 66 Fed. Reg. 1099 (2001), provide that "[t]he description need only describe in detail, that which is new or not conventional." 66 Fed. Reg. at 1106, citing *Hybritech v. Monoclonal Antibodies* 231 USPQ 81 (Fed. Cir. 1986), where the court held that "a patent need not teach, and preferably omits, what is well known in the art." *Id.* at 94. Accordingly, the present claims meet the requirement for written description set forth in the *Guidelines* and the supporting case law.

Claims 1-12 have also been rejected under 35 U.S.C. § 112, first paragraph, on the grounds that an insufficient written description is provided for the proteases recited in the claims. Claims 4 and 5 have been cancelled to expedite prosecution, rendering the rejection of this claim moot. It is respectfully submitted that the rejection should not be applied to claims 1-3 and 6-12 as amended or to new claims 38-47 for the reasons described below.

Claim 1 as amended recites a plant that is genetically modified to reduce or eliminate the activity of one or more proteases in its protein storage tissue, wherein said protease is selected from the group consisting of α-vacuolar processing enzyme, β-vacuolar processing enzyme, γ-vacuolar processing enzyme, ε-vacuolar processing enzyme, aspartic protease AP1, and aspartic protease AP2. The sequences of *Arabidopsis* α-VPE, β-VPE, γ-VPE, ϵ-VPE, AP1, and AP2 proteases are described in the specification on lines 4-13 of page 7. Furthermore, the domains of these proteins that are important for their proteolytic activity are known.

For example, vacuolar processing enzymes are members of the Peptidase C13 family of proteases and contain a conserved domain associated with protease activity. Appendix A shows the domain structure for the 11 proteins used to create the Pfam seed alignment for the Peptidase C13 protease family. Both α-VPE and β-VPE from *Arabidopsis*, as well as a VPE from *Citrus sinensis* (P49043) and *Ricinus communis* (P49042) are among the proteins used to create this seed alignment. Appendix B shows the sequence alignment for the proteins used to create the Pfam consensus alignment. It is apparent that the peptidase domain is highly conserved among the protein family members. Appendix C shows an alignment of ϵ-VPE (SEQ ID NO:2) with the Pfam Peptidase C13 family consensus alignment. This alignment demonstrates that ϵ-VPE

contains a peptidase C13 protease domain. Furthermore, the amino acids required for activity of the legumain family of proteases has been identified. *See*, Chen *et al.* (1998) *FEBS Letters* 441:361-365, provided herewith for the convenience of the Examiner as Appendix D.

The aspartic proteases AP1 and AP2 are members of the eukaryotic aspartyl protease family. Appendices E and F show alignments between the Pfam consensus sequence for the eukaryotic aspartyl protease family and AP1 (Appendix E) and AP2 (Appendix F) from *Arabidopsis*. Both proteins contain the Pfam consensus domain characteristic of with aspartyl protease activity.

The *Guidelines for Examination of Patent Applications Under the 35 U.S.C. § 112, ¶ 1, "Written Description Requirement,"* provide that a genus may be described by "sufficient description of a representative number of species . . . or by disclosure of relevant, identifying characteristics , *i.e.* structure or other physical and/or chemical properties." 66 Fed. Reg. 1099, 1106 (2001). Claims 1-3, 6-12, and 38-47 satisfy this requirement, because the specification provides examples of proteases that fall within the claimed genus of proteins, and the structural properties underlying the activity of these proteases are known.

In view of the above arguments and amendments, all grounds for rejection under the 35 U.S.C. § 112, first paragraph written description requirement have been obviated or overcome. Accordingly, reconsideration and withdrawal of the rejections are respectfully requested.

Claims 1-12 have further been rejected on the grounds that the specification does not provide sufficient enablement to allow one of skill in the art to make and use the claimed invention. Claims 4 and 5 have been cancelled, rendering the rejection of these claims moot. It is respectfully submitted that the rejection should not be applied to claims 1-3 and 6-12 as amended, or to new claims 38-47, for the reasons described below.

The Examiner argues that the specification does not provide methods to reduce the activity of an enzyme. However, methods of genetically modifying plants to reduce or eliminate the activity of a protease are described in the specification on lines 15 of page 7 through line 30 of page 8, line 16 of page 10 through line 3 of page 11, and line 29 of page 22 through line 11 of page 23. The methods of genetic modification include insertional mutagenesis (i.e. TUSC,

transposon tagging, T-DNA insertion), gene targeting (i.e. chimerplasty), sense suppression (co-suppression), antisense suppression, expression of specific protein processing protease inhibitors, or an increase in the level or activity of endogenous protease inhibitors. Accordingly, guidance for the reduction or elimination of protease activity is provided. Furthermore, working examples of plants containing an insertional event in β-VPE and ϵ-VPE are provided. *See*, pages 32-35 of the specification.

The Examiner also argues that the claims encompass antisense technology, and that antisense technology is unpredictable. Bryant (*et al.*) *Trends in Biotechnology* 7:20-21 is cited to support this argument. Bryant *et al.* teach that there have been increased reports that expression of specific genes can be down regulated by the presence of antisense RNA. *See*, first paragraph of Bryant *et al.* Bryant *et al.* also teach that in plants transformed to express an RNA antisense to chalcone synthase, the level of this enzyme varied from plant to plant. In some plants, chalcone synthase was extensively suppressed. *See,* column 3 of page 20 of Chalcone *et al.* The authors state that the reason for the variability in the levels of suppression of the enzyme is not known, but speculate that it may be due to a "position effect," in which the location of the transgene affects the expression level of the antisense RNA. The authors then conclude that antisense technology "raises the very real possibility of using recombinant DNA techniques to remove unwanted characteristics from economically important plant species." *See*, last paragraph of Bryant *et al.*

Accordingly, the major teaching of the Bryant *et al.* reference is that antisense technology holds great promise for the targeted reduction of gene expression in plants. Although not every plant transformed with an antisense construct in this reference showed extensive suppression of the target gene, plants having greater levels of target gene suppression could be identified by screening a population of transformed plants. Accordingly, Bryant *et al.* support the enablement of the present invention.

The Examiner also cites Martienssen (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:2021-2026 in support of the argument that essential genes cannot be down-regulated because suppression would lead to dominant lethal phenotypes that cannot be maintained. However, the present invention provides *working examples* of plants that are null for β-VPE or ϵ-VPE. *See*, pages 32-35 of the specification, demonstrating that down regulation of these enzymes is not lethal.

Accordingly, the teachings of the Martienssen reference are inapplicable to the enablement of the present invention.

Furthermore, as noted above, the methods of the invention encompass a number of methods for reducing or eliminating the activity of a plant protease. For example, the specification describes methods of reducing or eliminating the activity of a vacuolar processing enzyme by inhibiting its expression by sense suppression (cosuppression). See, line 30 of page 22 through line 11 of page 23. Cosuppresion has been successfully used to inhibit the expression of a number of plant genes. See, for example, U.S. Patent Nos. 5,283,184 and 5,034,323, cited on line 10 of page 23 of the specification. See also, U.S. Patent No. 5,942,657, which describes the successful use of antisense suppression and sense suppression to inhibit the expression of endogenous plant genes. These patents are provided herewith for the convenience of the Examiner as Appendices G, H, and I, respectively.

The Examiner also states that claim 6 is not enabled. because it encompasses methods of suppressing protease activity using fragments of SEQ ID NO:1 and cites Monan *et al.* (2002) *J. Virology* 76:1339-1348 for its teaching that "sugarcane plants expressing untranslated viral capsid sequences of *Sorghum mosaic virus* strain SCH, challenged with SrMV viruses of strains SCM and SCI and *Sugarcane mosaic virus* strain, show various levels of virus resistance that correlated with the percentage of sequence identity of the transgenes to the sequence of the challenging virus." Claim 6 has been canceled to expedite prosecution, rendering the rejection of this claim moot. In addition, the statement in Monan *et al.* refers to data shown in Ingelbrecht *et al.* (1999) *Plant Physiology* 119:1187-1197, provided herewith for the convenience of the Examiner as Appendix J. Ingelbrecht *et al.* teach that the greater the level of sequence similarity between the transgene and the challenging virus, the greater percentage of resistant plants. However, Ingelbrecht *et al.* show that approximately two thirds of plants transformed with a transgene having 95% sequence similarity with the challenging virus show resistance, and that some plants show resistance even when transformed with a transgene having only 75% identity with the challenging virus. See, Table III on page 1192 of Ingelbrecht *et al.* Accordingly, 100% sequence identity between the transgene and the challenging virus is not required to achieve resistance.

An enabling disclosure must describe the claimed invention in such a way as to enable the ordinarily skilled artisan to make and use the invention, and this description must be commensurate with the scope of the claimed invention. The test of enablement is not whether experimentation is necessary, but rather if experimentation *is* necessary, whether it is undue. *In re Angstadt*, 198 USPQ 214, 219 (C.C.P.A. 1976). The test of whether an invention requires undue experimentation is not based on a single factor, but rather a conclusion reached by weighing many factors. *In re Wands*, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988). Factors to be considered in determining whether undue experimentation is required include the quantity of experimentation necessary, the amount of guidance provided in the specification, the presence of working examples of the invention in the application, the nature of the invention, the state of the prior art, the relative skill of those in the art, the predictability in the art, and the breadth of the claimed invention. 8 USPQ2d at 1404. Accordingly, the holding of *Wands* does not require that an applicant identify every functional ϵ-VPE fragment or variant. Rather, *Wands* sets out factors to be considered in determining whether undue experimentation is required to make and use the claimed functional variants.

When all of the *Wands* factors are considered together, it is clear that although some quantity of experimentation may be required to reduce or eliminate the activity plant α-VPE, β-VPE, γ-VPE, ϵ-VPE, AP1, and AP2, the experimentation would not be undue in view of the nature of the invention, the state of the prior art (where the inhibition of the expression of endogenous plant genes is routine), the relative skill of those in the art, the predictability in the art (where a number of examples of suppression of endogenous plant genes have been described), the amount of direction provided in the specification (which provides methods for reducing or eliminating the activity of a plant protease and guidance regarding the proteases whose activity is to be reduced), the breadth of the claimed invention (which is limited to the reduction of the activity of a protease selected from α-VPE, β-VPE, γ-VPE, ϵ-VPE, AP1, and AP2), and the existence of working examples of plants having reduced protease activity. These factors all favor a conclusion that one of skill in the art could practice the claimed invention without undue experimentation.

In view of the above arguments and amendments, all grounds for rejection under 35 U.S.C. § 112, first paragraph, have been overcome. Reconsideration and withdrawal of the rejection are therefore respectfully requested.

The Rejection Under 35 U.S.C. § 102 Should be Withdrawn

Claims 1-3, 6-7, and 11-12 have been rejected under 35 U.S.C. § 102(b) on the grounds that they are anticipated by Hilder *et al.* (1987) *Nature* 300:160-63. It is respectfully submitted that the rejection should not be applied to the claims as amended, or to new claims 38-46 for the reasons described below.

Hilder *et al*. teach tobacco plants that have been genetically modified to comprise a cowpea trypsin inhibitor. Claim 1 as amended recites a plant that is genetically modified to reduce or eliminate the activity of one or more proteases in its protein storage tissue, wherein said protease is selected from the group consisting of α-vacuolar processing enzyme, β-vacuolar processing enzyme, γ-vacuolar processing enzyme, ε-vacuolar processing enzyme, aspartic protease AP1, and aspartic protease AP2. Accordingly, Hilder *et al.* do not anticipate claim 1 or dependent claims 2, 3, 6, 7, 11, 12, and 38-47.

In view of the above arguments and amendments, all grounds for rejection under 35 U.S.C. § 102, have been overcome. Reconsideration and withdrawal of the rejection are therefore respectfully requested.

The Rejection Under 35 U.S.C. § 103 Should be Withdrawn

Claims 1, 4, and 6-12 have been rejected under 35 U.S.C. § 103(a) on the grounds that they are obvious in view of U.S. Patent Nos. 5,955,653 and 5,723,763. Claim 4 has been cancelled to expedite prosecution, rendering the rejection of this claim moot. It is respectfully submitted that the rejection should not be applied to claims 1 and 6-12 as amended or to new claims 38-47 for the reasons described below.

U.S. Patent No. 5,955,653 teaches the use of the gene encoding the protease actinidin as a means of creating male-sterile plants. This patent does not suggest the inhibition of a protease.

U.S. Patent No. 5,723,763 teaches the use of antisense RNA to produce male-sterility and fertility-restored monocot plants. However, this patent does not teach the use of antisense RNA to inhibit the activity of a plant protease.

Claim 1 as amended recites a plant that is genetically modified to reduce or eliminate the activity of one or more proteases in its protein storage tissue, wherein said protease is selected from the group consisting of α-vacuolar processing enzyme, β-vacuolar processing enzyme, γ-vacuolar processing enzyme, ε-vacuolar processing enzyme, aspartic protease AP1, and aspartic protease AP2. Thus, U.S. Patent Nos. 5,955,653 and 5,723,763, taken individually or together do not teach or suggest the invention recited in claim 1 or its dependent claims.

In view of the above arguments and amendments, all grounds for rejection under 35 U.S.C. § 103, have been overcome. Reconsideration and withdrawal of the rejection are therefore respectfully requested.

CONCLUSIONS

It is believed that all the rejections have been obviated or overcome and the claims are in condition for allowance. Early notice to this effect is solicited. If in the opinion of the Examiner, a telephone conference would expedite the prosecution of the subject Application, the Examiner is invited to call the undersigned.

It is not believed that extensions of time or fees for net addition of claims are required, beyond those that may otherwise be provided for in documents accompanying this paper. However, in the event that additional extensions of time are necessary to allow consideration of this paper, such extensions are hereby petitioned under 37 CFR § 1.136(a), and any fee required therefore (including fees for net addition of claims) is hereby authorized to be charged to Deposit Account No. 16-0605.

Respectfully submitted,

Kathryn L. Coulter

Kathryn L. Coulter
Registration No. 45,889

CUSTOMER NO. 00826
ALSTON & BIRD LLP
Bank of America Plaza
101 South Tryon Street, Suite 4000
Charlotte, NC 28280-4000
Tel Raleigh Office (919) 862-2200
Fax Raleigh Office (919) 862-2260

CERTIFICATE OF EXPRESS MAILING

"Express Mail" mailing label number EV184328864US
Date of Deposit: December 17, 2003

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450

Nora C. Martinez

Nora C. Martinez

Pfam 10.0 (Saint L uis)
Home | Analyze a sequence | Browse Pfam | Keyword search | Taxonomy search | Swisspfam | Help

Domain structure f proteins in the Peptidase_C13 Seed alignment

Pfam and other annotation is shown as graphic objects along the sequence. One type of annotation may obscure other types; you can change the priority using the menu below. Priority is from left to right. You can also choose the blank menu entry to remove a type of annotation entirely. Mouse overs show start and end points and some additional information. Mouse clicks link to detailed descriptions or more information. More detailed help is available here.

signal peptide | pfamA | SMART | transmembrane | low complexity
coiled coil | pfamB | change order

GPI8_HUMAN Q92643 *GPI-anchor transamidase (EC 3.-.-.-) (GPI transamidase) (Phosphatidylinositol-glycan biosynthesis, class K protein)* [395 residues]

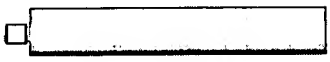
GPI8_YEAST P49018 *GPI-anchor transamidase (EC 3.-.-.-) (GPI transamidase). GPI-anchor transamidase (EC 3.-.-.-) (GPI transamidase).* [411 residues]

HGLB_SCHJA P42665 *Hemoglobinase precursor (EC 3.4.22.34) (Antigen Sj32). Hemoglobinase precursor (EC 3.4.22.34) (Antigen Sj32).* [423 residues]

HGLB_SCHMA P09841 *Hemoglobinase precursor (EC 3.4.22.34) (Antigen SM32). Hemoglobinase precursor (EC 3.4.22.34) (Antigen SM32).* [429 residues]

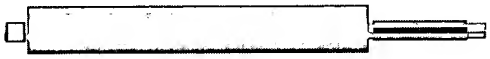
LEGU_CANEN P49046 *Legumain precursor (EC 3.4.22.34) (Asparaginyl endopeptidase). Legumain precursor (EC 3.4.22.34) (Asparaginyl endope* [475 residues]

O46047 O46047 *EG:133E12.3 protein (CG4406 protein). EG:133E12.3 protein (CG4406 protein).* [326 residues]



VPEA_ARATH P49047 *Vacuolar processing enzyme, alpha-isozyme precursor (EC 3.4.22.-) (Alpha-VPE). Vacuolar processing enzyme, alpha-iso* [478 residues]

VPEB_ARATH Q39044 *Vacuolar processing enzyme, beta-isozyme precursor (EC 3.4.22.-) (Beta-VPE). Vacuolar processing enzyme, beta-isozym* [486 residues]



VPE_CITSI P49043 *Vacuolar processing enzyme precursor (EC 3.4.22.-) (VPE). Vacuolar processing enzyme precursor (EC 3.4.22.-) (VPE).* [494 residues]

VPE_RICCO P49042 *Vacuolar processing enzyme precursor (EC 3.4.22.-) (VPE). Vacuolar processing enzyme precursor (EC 3.4.22.-) (VPE).* [497 residues]

Home | Analyze a sequence | Browse Pfam | Keyword search | Taxonomy search | Swisspfam | Help
Comments, questions, flames? Email <pfam@genetics.wustl.edu>.

```
# STOCKHOLM 1.0
#=GF ID    Peptidase_C13
#=GF AC    PF01650
#=GF DE    Peptidase C13 family
#=GF GA    -94.0 -94.0
#=GF NC    -154.0 -154.0
#=GF TC    -31.9 -31.9

VPE_VICSA/10-374   LFFTIVVTFLTVVSSGRDL...PGDYLRLPSETSRFFREP......KNDD
VPE_CITSI/8-375    VLITLLVALAGIADGSRDI...AGDILKLPSEAYRFFHNGGGGAKVNDDD
VPEA_ARATH/5-360   VSFLALFLFLVAAVSG........DVIKLPSLASKFFRPT......ENDD
VPE_RICCO/17-378   LVFTLSFLPIPGLLASRLN..PFEPGILMPTEEAEPVQV........DDD
LEGU_CANEN/1-356   MVMMLVMLSLHG.TAARLNRREWDSVIQLPTEPVD..............
VPEB_ARATH/10-367  ALLLLLVLLVHA..ESRGR...FEPKILMPTEEANPAD.........QDE
HGLB_SCHMA/5-332   SLFLISILHILLVKCQL......DTNYEVSDETVS.............D
HGLB_SCHJA/4-326   SIFFIHILRIVLVDCN..........EYSEENVD.............D
O46047/9-326       FVVIFALILASCRVEA........DNTSVLPEGFVDAA..........QR
GPI8_HUMAN/12-325  TVLATVLLLSFGSVAA........SHIEDQAEQFF.............R
GPI8_YEAST/9-337   LLLLYIFLLPLS..GA........NNTDAAHEVIA..............

VPE_VICSA/10-374   DFEGTRWAILLAGSNGYWNYRHQSDVCHAYQLLRKGGSKEENIIVFMYDD
VPE_CITSI/8-375    DSVGTRWAVLLAGSNGFWNYRHQADICHAYQLLRKGGLKDENIIVFMYDD
VPEA_ARATH/5-360   DS..TKWAVLVAGSSGYWNYRHQADVCHAYQLLKKGGVKEENIVVFMYDD
VPE_RICCO/17-378   DQLGTRWAVLVAGSMGFGNYRHQADVCHAYQLLRKGGLKEENIIVFMYDD
LEGU_CANEN/1-356   DEVGTRWAVLVAGSNGYGNYRHQADVCHAYQLLIKGGVKEENIVVFMYDD
VPEB_ARATH/10-367  DGVGTRWAVLVAGSSGYGNYRHQADVCHAYQILRKGGLKEENIVVLMYDD
HGLB_SCHMA/5-332   NN...KWAVLVAGSNGYPNYRHQADVCHAYHVLRSKGIKPEHIITMMYDD
HGLB_SCHJA/4-326   RH...KWAVLVAGSNGFENYRHQADVCHAYHVLLSKGVKPEHIITFMYDD
O46047/9-326       STHTNNWAVLVDASRFWFNYRHVANVLSIYRSVKRLGIPDSQIILMIADD
GPI8_HUMAN/12-325  SGHTNNWAVLVCTSRFWFNYRHVANTLSVYRSVKRLGIPDSHIVLMLADD
GPI8_YEAST/9-337   .TNTNNWAVLVSTSRFWFNYRHMANVLSMYRTVKRLGIPDSQIILMLSDD

VPE_VICSA/10-374   IASNEENPRPGVIINKP.DGDDVYAG.VPKDYTGAEVHADNFYAALLGNK
VPE_CITSI/8-375    IAFNEENPRPGVIINHP.HGDDVYKG.VPKDYTGEDVTVEKFFAVVLGNK
VPEA_ARATH/5-360   IAKNEENPRPGVIINSP.NGEDVYNG.VPKDYTGDEVNVDNLLAVILGNK
VPE_RICCO/17-378   IAKNELNPRPGVIINHP.QGEDVYAG.VPKDYTGEHVTAKNLYAVLLGDK
LEGU_CANEN/1-356   IAYNAMNPRPGVIINHP.QGPDVYAG.VPKDYTGEDVTPENLYAVILGDK
VPEB_ARATH/10-367  IANHPLNPRPGTLINHP.DGDDVYAG.VPKDYTGSSVTAANFYAVLLGDQ
HGLB_SCHMA/5-332   IAYNLMNPFPGKLFNDY.NHKDWYEG.VVIDYRGKNVNSKTFLKVLKGDK
HGLB_SCHJA/4-326   IAHNKENPFPGKIFNDY.RHKDYYKG.VVIDYKGKKVNPKTFLQVLKGDK
O46047/9-326       MACNARNPRPGQVYNNANQHINVYGDDVEVDYRGYEVTVENFVRLLTGRT
GPI8_HUMAN/12-325  MACNPRNPKPATVFSHKNMELNVYGDDVEVDYRSYEVTVENFLRVLTGRI
GPI8_YEAST/9-337   VACNSRNLFPGSVFNNKDHAIDLYGDSVEVDYRGYEVTVENFIRLLTDRW

VPE_VICSA/10-374   SALTGGSGKVVDSGPNDHIFVYYTDHGGPGVLGMPVGPYLYASDLNEVLK
VPE_CITSI/8-375    TALTGGSGKVVDSGPNDHIFIFYSDHGGPGVLGMPTSRYIYADELIDVLK
VPEA_ARATH/5-360   TALKGGSGKVVDSGPNDHIFIYYSDHGGPGVLGMPTSPNLYANDLNDVLK
VPE_RICCO/17-378   SAVQGGSGKVVDSKPNDRIFLYYSDHGGPGVLGMPNLPYLYAMDFIEVLK
LEGU_CANEN/1-356   SKVKGGSGKVINSNPEDRIFIFYSDHGGPGVLGMPNAPFVYAMDFIDVLK
VPEB_ARATH/10-367  KAVKGGSGKVIASKPNDHIFVYYADHGGPGVLGMPNTPHIYAADFIETLK
HGLB_SCHMA/5-332   S...AG.GKVLKSGKNDDVFIYFTDHGAPGLIAFPDD.ELYAKEFMSTLK
HGLB_SCHJA/4-326   R...AG.GKVLKSGKNDDVFIYFTDHGAPGILAFPDD.DLHAKPFINTLK
O46047/9-326       QNGTAR.SKKLLSDAGSNVLIYLTGHGGDGFLKFQDSEEITSQELADGIQ
GPI8_HUMAN/12-325  PPSTPR.SKRLLSDDRSNILIYMTGHGGNGFLKFQDSEEITNIELADAFE
GPI8_YEAST/9-337   TEDHPK.SKRLLTDENSNIFIYMTGHGGDDFLKFQDAEEIASEDIADAFQ

VPE_VICSA/10-374   KKHASGTYKSLVFYLEACESGSIFEGLLPDDLNIYATTASNAEESSWGYY
VPE_CITSI/8-375    KKHASGNYKSLVFYLEACESGSIFEGLLLEGLNIYATTASNAEESSWGTY
VPEA_ARATH/5-360   KKYASGTYKSLVFYLEACESGSIFEGLLPEGLNIYATTASNAEESSWGTY
VPE_RICCO/17-378   KKHAAGGYKKMVIYVEACESGSIFEGIMPKDVDIYVTTASNAQESSWGTY
```

```
LEGU_CANEN/1-356    KKHASGGYKEMVIYIEACESGSIFEGIMPKDLNIYVTTASNAQENSFGTY
VPEB_ARATH/10-367   KKHASGTYKEMVIYVEACESGSIFEGIMPKDLNIYVTTASNAQESSYGTY
HGLB_SCHMA/5-332    YLHSHKRYSKLVIYIEANESGSMFQQILPSNLSIYATTAANSTECSYSTF
HGLB_SCHJA/4-326    YLRQHRRYSKLVIYVEACESGSMFAGLLPTDINIYATTAARPDESSYATF
O46047/9-326        QMWEKKRYNELFFMVDTCQAASLYEKFTSP..NVLAVASSLVGEDSLSHH
GPI8_HUMAN/12-325   QMWQKRRYNELLFIIDTCQGASMYERFYSP..NIMALASSQVGEDSLSHQ
GPI8_YEAST/9-337    QMYEKKRYNEIFFMIDTCQANTMYSKFYSP..NILAVGSSEMDESSYSHH

VPE_VICSA/10-374    CPGDKPPPPPEYSTCLGDLYSIAWMEDSEVHNLQTESLQQQYKLVKNRTI
VPE_CITSI/8-375     CPGEIPGPPPEYSTCLGDLYSIAWMEDSDIHNLRTETLHQQYELVKTRTA
VPEA_ARATH/5-360    CPGEDPSPPSEYETCLGDLYSVAWIEDSEKHNLQTETLHEQYELVKKRTA
VPE_RICCO/17-378    CPGMEPSPPPEFTTCLGDLYSVAWMEDSESHNLKKETVKQQYSSVKARTS
LEGU_CANEN/1-356    CPGMNPPPPEEYVTCLGDLYSVSWMEDSETHNLKRETVQQQYQSVRKRTS
VPEB_ARATH/10-367   CPGMNPSPPSEYITCLGDLYSVAWMEDSETHNLKKETIKQQYHTVKMRTS
HGLB_SCHMA/5-332    CG..DP....TITTCLADLYSYNWIVDSQTHHLTQRTLDQQYKEVKRETD
HGLB_SCHJA/4-326    CD..DP....RISSCLADLYSYDWIVDSEKHQLTQRTLDQQYKEVKFETN
O46047/9-326        VDPSIG....VYMIDRYTYYALEFLEKVQP..FSKRTIGEFLQVCPKRVC
GPI8_HUMAN/12-325   PDPAIG....VHLMDRYTFYVLEFLEEINP..ASQTNMNDLFQVCPKSLC
GPI8_YEAST/9-337    SDVEIG....VAVIDRFTYYCLDFLEQIDK..NSTLTLQDLFDSFTFEKI

VPE_VICSA/10-374    S...EPYGSHVMEYGDIGLSKNDLYQYLGTNPANDNNSFVDETENSLKLR
VPE_CITSI/8-375     SY..NSYGSHVMQYGDIGLSKNNLFTYLGTNPANDNYTFVDENS....LR
VPEA_ARATH/5-360    G.SGKSYGSHVMEFGDIGLSKEKLVLFMGTNPADENFTFVNENS....IR
VPE_RICCO/17-378    NYNTYAAGSHVMQYGNQSIKADKLYLFQGFDPASVNFPPNNAHL.....N
LEGU_CANEN/1-356    NSNSYRFGSHVMQYGDTNITAEKLYLYHGFDPATVNFPPHNGNL.....E
VPEB_ARATH/10-367   NYNTYSGGSHVMEYGNNSIKSEKLYLYQGFDPATVN.LPLNELP....VK
HGLB_SCHMA/5-332    L.......SHVQRYGDTRMGKLYVSEFQGSRDKSS..SENDEPP.....M
HGLB_SCHJA/4-326    L.......SHVQRYGDKKMGKLYLSEFQGSRKKAS..TEHDEPP.....M
O46047/9-326        I.......STVGVRKDLYPRDPHKVPITDFFGAIR..............
GPI8_HUMAN/12-325   V.......STPGHRTDLFQRDPKNVLITDFFGSVR..............
GPI8_YEAST/9-337    H.......SHVGVRTDLFDRNPSEVLITDFFANVQN.VIPDDSK......

VPE_VICSA/10-374    TPSAAVNQRDADLIHFWEKFRKAPEGSSQ
VPE_CITSI/8-375     PASKAVNQRDADLLHFWDKYRKAPEGTPR
VPEA_ARATH/5-360    PPSRVTNQRDADLVHFWHKYQKAPEGSAR
VPE_RICCO/17-378    APMEVVNQRDAELHFMWQLYKRSENGSEK
LEGU_CANEN/1-356    AKMEVVNQRDAELLFMWQMYQRSNHQPEK
VPEB_ARATH/10-367   SKIGVVNQRDADLLFLWHMYRTSEDGSRK
HGLB_SCHMA/5-332    KPRHSIASRDIPLHTLHRQIMMTNN.AED
HGLB_SCHJA/4-326    KPKDSIPSRDIPLHTLHRRIMMANN.MND
O46047/9-326        .PTRVS........TDRINVTLAN...EE
GPI8_HUMAN/12-325   .KVEIT........TETIKLQQDS...EI
GPI8_YEAST/9-337    .PLSVSHYHHYKDHIDTAQYELNNNVLDL
//
```

Pfam 10.0 (Saint L uis)
Home | Analyze a sequence | Browse Pfam | Keyword search | Taxonomy search | Swisspfam | Help

Starting search. Estimated time: 14 seconds (assuming all Wulfpack nodes are running). Please wait...

Pfam HMM search results, glocal+local alignments merged (Pfam_ls+Pfam_fs)

[Go here for an explanation of the format of the results]

Model	Seq-from	Seq-to	HMM-from	HMM-to	Score	E-value	Alignment	Description
!! Peptidase_C13	11	362	1	362	599.2	**2.7e-177**	glocal	Peptidase C13 family

Alignments of top-scoring domains:

Format for alignment of query to Seed: Plain text

```
Peptidase_C13: domain 1 of 1, from 11 to 362: score 599.2, E = 2.7e-177
                  *->lvlllvllllllvvaardnfddnielpsEevdffrdddtvgtrWAVL
                     lv+l +ll  ++  ++      +++l   +     d + +gtrWAVL
E-VPE query    11    LVFLHALLIFSA--ES-----RKTQLLNDNDVESSDKSAKGTRWAVL 50

                  VaGSnGyfNYRHqADVchaYqlLkkgGiKdenIIvfmyDDiAcNerNPrP
                  VaGSn y NYRHqAD+chaYq+L+kgG+KdenIIvfmyDDiA+ ++NPrP
       query    51 VAGSNEYYNYRHQADICHAYQILRKGGLKDENIIVFMYDDIAFSSENPRP 100

                  GviiNhpnqgkDvYaGdVpvDYrGeeVtvenFlaVLlGdksaltggSGKv
                  GviiN+p +g+DvY+G Vp+DY+ e V+v nF+ VLlG+ s +tgg GKv
       query   101 GVIINKP-DGEDVYKG-VPKDYTKEAVNVQNFYNVLLGNESGVTGGNGKV 148

                  ldSgpnDnIFIYytDHGGpGvLgfPdseeiyAkdlidvLkkkhasgrYke
                  ++SgpnDnIFIYy+DHG+pG++++P+  e+ Akd+++vL+k+h  + Y++
       query   149 VKSGPNDNIFIYYADHGAPGLIAMPTGDEVMAKDFNEVLEKMHKRKKYNK 198

                  LvfyiEACeSgSiFegllpkdLNIyAtTASnaeEsSygtycdgeiPsPPp
                  +v+y+EACeSgS+Feg+l k+LNIyA+TA+n+ EsS+g yc+ + P+PP+
       query   199 MVIYVEACESGSMFEGILKKNLNIYAVTAANSKESSWGVYCPESYPPPPS 248

                  eyvtCLgDLYslaWlEdsekHnlskeTlqqqYksvkkrtslyntysyGSH
                  e+ tCLgD  s++WlEds+ H +skeTl+qqY++vk+r+ +    +  SH
       query   249 EIGTCLGDTFSISWLEDSDLHDMSKETLEQQYHVVKRRVGS--DVPETSH 296

                  VmeygDlglrkeklvlftGffpavrNftfvdenplrkpse...vvnqRDa
                  V ++g  ++ k++l+ + G +p ++Nftf +  +  +p+++++ vn RD+
```

```
query   297 VCRFGTEKMLKDYLSSYIGRNPENDNFTFTESFS--SPISnsgLVNPRDI 344

            dLhtlwrkyqkanngsek<-*
            +L+ l rk+qka+ gs +
query   345 PLLYLQRKIQKAPMGSLE    362
```

align query/11-362 to Peptidase_C13(ls) Seed

NEW! Phylogenomic analysis of query using RIO.

Given a query sequence, Pfam domain, and species, the RIO server will order sequences in the Pfam domain by orthology to the query. Many other options are available, and an annotated gene tree can be generated and viewed with ATV. The button below will send your query and Pfam domain hits to the RIO server.

send query to RIO

Home | Analyze a sequence | Browse Pfam | Keyword search | Taxonomy search | Swisspfam | Help

Comments, questions, flames? Email <pfam@genetics.wustl.edu>.

FEBS 21321 FEBS Letters 441 (1998) 361–365

Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases

Jinq-May Chen, Neil D. Rawlings, Richard A.E. Stevens, Alan J. Barrett*

MRC Molecular Enzymology Laboratory, Babraham Institute, Babraham, Cambridge CB2 4AT, UK

Received 10 November 1998

Abstract We show by site-directed mutagenesis that the catalytic residues of mammalian legumain, a recently discovered lysosomal asparaginycysteine endopeptidase, form a catalytic dyad in the motif His-Gly-spacer-Ala-Cys. We note that the same motif is present in the caspases, aspartate-specific endopeptidases central to the process of apoptosis in animal cells, and also in the families of clostripain and gingipain which are arginyl/lysyl endopeptidases of pathogenic bacteria. We propose that the four families have similar protein folds, are evolutionarily related in clan CD, and have common characteristics including substrate specificities dominated by the interactions of the S1 subsite.

Key words: Legumain; Caspase; Clostripain; Gingipain; Peptidase clan

1. Introduction

Legumain (EC 3.4.22.34) is a cysteine endopeptidase that was until recently known only from plants [1,2] and *Schistosoma* [3]. In plants there is abundant evidence that legumain performs a protein-processing function, causing limited proteolysis of precursor proteins and protein splicing [2,4]. Following the discovery of legumain in mammalian cells, and the cloning and sequencing of legumain from human [5] and mouse [6], we have shown the enzyme to be predominantly lysosomal [6], but it has strict specificity for the hydrolysis of bonds on the carboxyl side of asparagine which is very different from that of any other lysosomal endopeptidase and adapts it particularly for limited proteolysis. Consistent with this, there is evidence for a key role of legumain in the processing of a bacterial antigen for the MHC class II system in the lysosomal/endosomal system of antigen presenting cells [7] and it also processes progelatinase A to the active enzyme (J.-M. Chen and A.J. Barrett, unpublished results).

The amino acid sequence of legumain shows it to belong to a distinct family of cysteine endopeptidases that has been termed peptidase family C13. (This and other identifiers for peptidase families and clans are as defined in the MEROPS database [8,9].) We here describe the identification of the catalytic residues of legumain, and show how they reveal relationships between family C13 and several other families of cysteine endopeptidases.

*Corresponding author. Fax: (44) (1223) 496023.
E-mail: alan.barrett@bbsre.ac.uk

2. Materials and methods

2.1. Mutagenesis and plasmid construction of legumain mutants

Mouse legumain cDNA was removed from the plasmid construct pCMVmusleg [6] with *Eco*RI and *Xho*I and subcloned into the pCR-Script Amp^r SK(+) vector (Stratagene) at the same sites to produce a pCRmusleg construct. Mutagenesis was performed by PCR site-directed mutagenesis as described by Picard et al. [10] with modifications. Six mutagenic primers based on the nucleotide sequence of human legumain were designed to generate mutants H47A, C52S, H150A, H164A, C191S, and S195A. These primers with underlined nucleotides encoding the mutated residues were: 5'-GGTATAATTATAGGGCACAGGCAGACGC-3' for H47A, 5'-AGGCAGACGCGTCCCATGCCTAC-3' for C52S, 5'-CATTTACTTCACTGACGCTGGATCTACTGG-3' for H150A, 5'-CTAATGATGATCTTGCTGTCAAGGACCTGA-3' for H164A, 5'-CTACATTGAAGCCTCTGAGTCTGGGTCC-3' for C191S, and 5'-GTGAGTCTGGGGCTATGATGAACC-3' for S195A. Since the mutagenic primers were based on the human legumain sequence, some silent mutations were introduced into the mouse legumain constructs, and in mutant H150A the residue Ala-152 of mouse legumain was replaced by the Ser corresponding in the human sequence. Polymerase *Pfu* was used for all PCR reactions with a typical run of 95°C×1 min, 48°C×1 min, 72°C×2 min for 30 cycles. Introduction of the mutated codon was performed by PCR using a mutagenic primer, the downstream primer of mouse legumain (5'-CTGGTGTTGGTGTGGGACTTGACC-3'), and pCMVmusleg as the template to generate a megaprimer. The megaprimer and the legumain-specific upstream primer (5'-GACGCCCGGAATTCCCACGGTT-3') were used for the second run of PCR with the same template to generate an 880-bp product containing the mutated nucleotide(s). This product was purified from the agarose gel by use of the Gene Clean Kit (BIO 101) and double digested with *Eco*RI and *Hin*cII. Wild-type pCRmusleg was similarly digested with *Eco*RI and *Hin*cII to remove the 880-bp wild type sequence, and the mutated PCR product was ligated into this vector to generate the mutated forms of pCRmusleg. Plasmid DNA of wild-type and mutant pCRmusleg was propagated in *E. coli* DH10B (Life Technologies) or XL10-gold (Stratagene) and purified by standard procedures [11]. The sequences of all mutated constructs were confirmed by sequencing. The full length cDNA of mutant legumain was then excised from the cloning vector pCRmusleg with *Eco*RI and *Xho*I and cloned into the mammalian expression vector pCMV-SPORT2 for transient expression.

2.2. Cell culture and transfection

Human embryonic kidney (HEK) 293 cells (ATCC no. CRL 1573) were maintained in minimal Eagle's medium supplemented with 10% fetal bovine serum and antibiotics. The calcium phosphate precipitation method was used to transfect HEK 293 cells with the wild-type and mutant legumain constructs as described [6].

2.3. Detection of recombinant legumain and enzymatic assays

HEK 293 cells transfected with legumain constructs and mock-transfected cells were harvested three days after transfection. Cells were disrupted by three cycles of freeze and thaw in the lysis buffer (0.1 M sodium citrate, pH 5.8, containing 1 mM EDTA and 1% *n*-octyl β-D-glucopyranoside). Cell lysates were collected after centrifugation at 18000×*g* for 10 min. The recombinant protein of wild-type and mutant legumain expressed in the cell lysate were detected by SDS-PAGE immunoblot developed with a polyclonal antibody specific to pig legumain [6]. To measure the enzymatic activity of recombinant legumain, lysate (5 μl) was placed in the well of a 96-

PII: S0014-5793(98)01574-9

```
           3         4         5            15        16                  19
           012345678901234567890123456   34567890123456789012a345678   1234567890123456789
                                  *    *           *                *                   *   *
AJ000990   KHWVVIVAGSNGWYNYRHQADACHAY...VFIYFTDHGATGILVFPNDD-LHVKDL...YQKMVFYIEACESGSMMNH
Q99538     KHWVVIVAGSNGWYNYRHQADACHAY...VFIYFTDHGSTGILVFPNED-LHVKDL...YRKMVFYIEACESGSMMNH
Q92643     NNWAVLVCTSRFWFNYRHVANTLSVY...ILIYMTGHGGNGFLKFQDSEEITNIEL...YNELLFIIDTCQGASMYER
P49046     TRWAVLVAGSNGYGNYRHQADVCHAY...IFIFYSDHGGPGVLGMPNAPFVYAMDF...YKEMVIYIEACESGSIFEG
P49042     TRWAVLVAGSM FGNYRHQADVCHAY...IFLYYSDHGGPGVLGMPNLPYLYAMDF...YKKMVIYVEACESGSIFEG
P09841     NKWAVLVAGSNGYPNYRHQADVCHAY...VFIYFTDHGAPGLIAFPDDE-LYAKEF...YSKLVIYIEANESGSMFQQ
P42665     HKWAVLVAGSNGFENYRHQADVCHAY...VFIYFTDHGAPGILAFPDDD-LHAKPF...YSKLVIYVEACESGSMFAG
```

Fig. 1. Selection of candidate catalytic residues for mutagenesis. Only the parts of the sequences containing well-conserved residues of cysteine or histidine are shown. The sequences (numbered according to mouse preprolegumain [6]) are: AJ000990, mouse legumain; Q99538, human legumain; Q92643, human GPI8 protein; P49046, jack bean (*Canavalia ensiformis*) asparaginyl endopeptidase; P49042, castor bean (*Canavalia ensiformis*) vacuolar processing enzyme; P09841, *Schistosoma mansoni* haemoglobinase; and P42665, *Schistosoma japonicum* haemoglobinase.

well microtiter plate and the reaction was started by addition of 120 μl of benzyloxycarbonyl-Ala-Ala-Asn-7-(4-methyl)coumarylamide (Z-Ala-Ala-Asn-NHMec) solution in assay buffer (39.5 mM citric acid, 121 mM Na_2HPO_4, pH 5.8, containing 1 mM DTT, 1 mM EDTA and 0.1% CHAPS) to give 10 μM final concentration of the substrate. The plates were incubated at 25°C and readings of fluorescence (excitation 360 nm, emission 460 nm) were taken at 10 and 130 min in a PerSeptive Biosystems CytoFluor II fluorescence plate reader. The change between the two readings was the measure of activity. Specific legumain activity was obtained by subtracting the control value of mock-transfected cells.

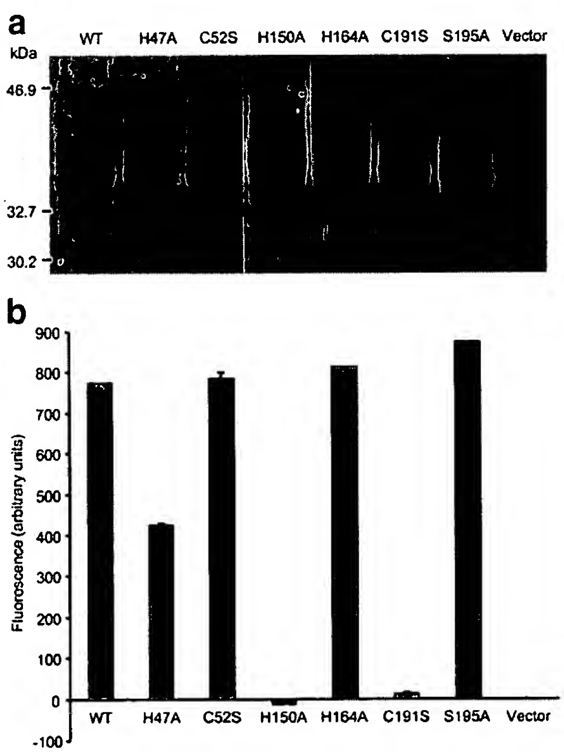


Fig. 2. Expression and activity of wild-type and mutant forms of legumain. a: Lysates (5 μl except H150A 20 μl) of HEK 293 cells transfected with legumain constructs in the pCMV-SPORT2 vector, or mock-transfected cells, were run in SDS-PAGE, and transblotted for the detection of legumain as described in Section 2. b: Legumain activity of cell lysates as in a, assayed with Z-Ala-Ala-Asn-NHMec as described in Section 2. The blank value of activity in the mock-transfected cells that has been subtracted is 622 fluorescence units. Bars show S.D. ($n = 3$).

3. Results and discussion

3.1. Catalytic dyad of legumain

The catalytic residues in all known cysteine endopeptidases are a dyad of cysteine and histidine, which may occur in either order, cysteine–histidine or histidine–cysteine, in the amino acid sequence. Accordingly, the residues His-47, Cys-52, His-150 and Cys-191 were selected as targets for mutagenesis by inspection of an alignment of the sequences for family C13, part of which is shown in Fig. 1. Site directed mutagenesis was done as described in Section 2 to generate mutant constructs H47A, C52S, H150A and C191S. Two additional constructs, H164A and S195A, were made as controls. Wild-type and mutant proteins were transiently expressed in HEK 293 cells under the control of the CMV promoter. Immunoblotting analysis showed that recombinant proteins from all constructs were expressed and processed to mature forms similar in molecular size (35 kDa) to the wild-type protein (Fig. 2a). The H150A mutant was expressed about four-fold less efficiently than the other forms as judged by inspection of the immunoblots.

Assays of legumain activity in extracts of cells expressing the mutated forms of the enzyme revealed that mutants H150A and C191S were completely inactive, whereas H47A and C52S were 54% and 101%, respectively, as active as the wild type. The two control mutants, H164A and S195A, were also as active as the wild type (Fig. 2b). In separate experiments (not shown) the C191S mutant of human legumain was made and again was found to be expressed normally but to be inactive. We conclude that His-150 and Cys-191 are essential for the catalytic activity of legumain and can be assumed to represent the catalytic dyad of the enzyme. The legumain of *Schistosoma mansoni* apparently lacks Cys-191 (Fig. 1), and we assume that the published sequence is that of an inactive variant of the enzyme. This would be consistent with the failure of Götz and Klinkert [12] to detect activity of the *Schistosoma mansoni* enzyme expressed in insect cells, but Cys-191 is present in the *S. japonicum* sequence. The other generally conserved cysteine residue, Cys-52, is absent from the human GPI8 protein that catalyses the attachment of glycosylphosphatidylinositol anchors to proteins [13,14].

3.2. An active site motif recognised in legumain is present in other families

The amino acid sequence of the segment of the polypeptide chain containing the catalytic residues identified in legumain

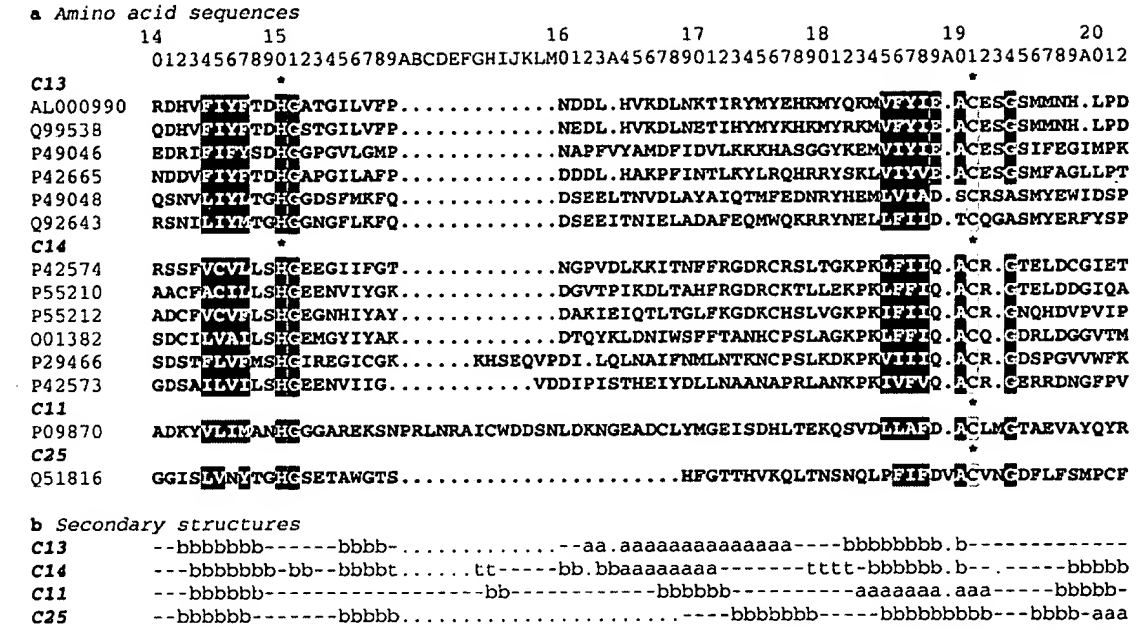


Fig. 3. Alignment of the catalytic sites of cysteine endopeptidases of families of legumain (C13), caspase-1 (C14), clostripain (C11) and gingipain R (C25). a: The sequence segments containing the known or putative catalytic residues (magenta and yellow) in each of the four families were aligned manually by the introduction of gap residues '.'. It can be seen that two or three residues N-terminal to each of the catalytic residues there is a block of four predominantly hydrophobic residues (green). Other residues that tend to be conserved between the families are printed in white on black. Asterisks are placed over the His and Cys residues in each family for which they have been identified experimentally as catalytic, and in addition the His residue in family C25 was proposed to be catalytic by Pavloff et al. [19] on the basis of its conservation in the family. Key to sequences: AL000990, mouse legumain; Q99538, human legumain; P49046, jack bean (*Canavalia ensiformis*) asparaginyl endopeptidase; P42665, *Schistosoma japonicum* haemoglobinase; P49048, *Caenorhabditis elegans* hypothetical protein T05E11.6; Q92643, human GPI8 protein; P42574, human caspase 3; P55210, human caspase 7; P55212, human caspase 6; O01382, *Drosophila melanogaster* caspase; P29466, human caspase 1; P42573, *Caenorhabditis elegans* CED3 protein; P09870, *Clostridium histolyticum* α-clostripain; Q51816, *Porphyromonas gingivalis* gingipain R. b: The corresponding secondary structures are those determined crystallographically for human caspase-1 (family C14; PDB 1IBC) or predicted for the C13, C11 and C25 families by the PHD program [20]. The β-strand residues are marked 'b', the α-helix residues 'a' and the turn residues 't'. Both catalytic residues in family C14 are preceded by β-strands (strands S3 and S4, respectively) and with one exception these strands (containing the blocks of hydrophobic residues noted in a above) are predicted for the other families. The secondary structure elements shown here for the caspase family can be seen in the three-dimensional representation of Fig. 4.

was compared with sequences of cysteine peptidases in other families. As shown in Fig. 3, it was found that a motif His-Gly-spacer-Ala-Cys could be recognised not only in the legumain family but also in the families of caspase-1 (C14), clostripain (C11) and gingipain R (C25). The caspases are mammalian cytosolic endopeptidases that play key roles in apoptosis, whereas clostripain and gingipain are cysteine endopeptidases from the pathogenic bacteria *Clostridium histolyticum* and *Porphyromonas gingivalis*, respectively [15,16]. The cysteine residues in this motif have already been identified as the catalytic cysteines of all three of these other families [17–19] and the histidine has been identified as catalytic in C14 [18] and suggested tentatively for C25 [19]. Closely preceding each of the catalytic residues in all four families is a block of four hydrophobic amino acids (marked in blue in Fig. 3a). The corresponding parts of the secondary structures predicted for mouse legumain, clostripain, and gingipain R by

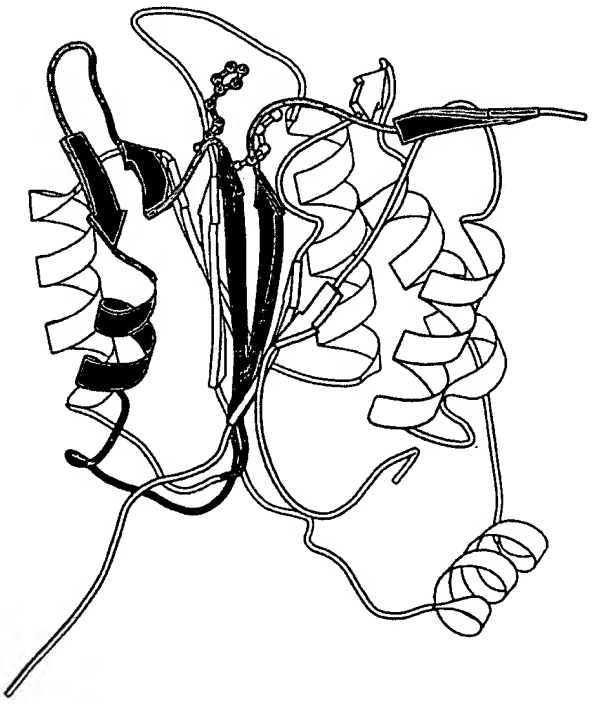


Fig. 4. Tertiary structure of part of the caspase-1 molecule (peptidase family C14) drawn from PDB record 1IBC by use of the MOLSCRIPT program [26]. Shown in green are the two β-strands that support the catalytic residues, histidine (pink) and cysteine (yellow), and the remainder of the segment included in the alignment of Fig. 3 is coloured in cyan.

the PHD program [20] are shown in Fig. 3b, and are generally in good agreement with the structure known for caspase-1 [18]. The longest spacer sequence exists in clostripain, but it is notable that the maturation of clostripain normally involves the autolytic excision of the nonapeptide 156-159E, although this is not required for activity [21]. As shown in Fig. 4, the catalytic residues of caspase-1 occur at the ends of two parallel β-strands. These are two central strands in a set of six that form a β-sheet at the core of the molecule, and the blocks of hydrophobic residues seen in the alignment are contained in the strands. This arrangement is expected to be mirrored in the other families.

3.3. The active site motif is unique

We next asked whether the sequence motif that contains the catalytic dyad in each of these four families of cysteine endopeptidases occurs in other proteins too. First, Motif 1 (Table 1) was formulated to include the blocks of hydrophobic residues preceding each of the catalytic residues as highlighted in Fig. 3, and the two non-hydrophobic residues immediately preceding the catalytic cysteine, and this was used to search the protein sequence databases. All 37 sequences retrieved proved to be known members of peptidase families C11, C13, C14 or C25, but many peptidases of these families were not recognised by Motif 1, so the more relaxed Motif 2 was constructed. This allowed a greater or lesser number of residues spacing the catalytic His and Cys, any hydrophobic residue at any position in the blocks preceding each member of the catalytic dyad, and restrictions on residue 189. With Motif 2, 66 hits were retrieved, and these included all but seven of the known members of the four families together with just three apparently unrelated false positive sequences all for hypothetical proteins (Table 1).

It is evident from these results that Motif 2 is very effective in screening protein sequences for the peptidases of just four families. Sequences belonging to all other known families of cysteine peptidases were quite unrecognised, notably including those from the papain (C1) and the calpain (C2) families, and those of many viruses.

Table 1
Database searches for proteins containing sequence motifs based on the active site sequence identified in legumain

Motif 1:	[AFILV]-[CILV]-[AFINVY]-[FILMY]-[ALMST]-X-H-G-X(2-4,47)-[ILPV]-[FILV]-[AFIY]-[AFIV]-X(1,2)-A-C
Motif 2:	[AFILMVWY]-[ACFILMVWY]-[ACFILMNVWY]-[AFILMVWY]-[AFILMSTVWY]-X-H-G-X(20,60)-[AFILMPVWY]-[AFILMVWY](3)-[EDQN]-X(0,1)-[ACST]-C

Sequences retrieved

C11	P09870
C13	O24325, O24326, O46047, P42665, P49018, P49042, P49043, P49044, P49045, P49046, P49047, P49048, Q39044, Q39119, Q92643, Q99538
C14	O01382, O02002, O02229, O02433, O08736, O08738, O14676, O14822, O35397, O35669, O42284, O55194, P29452, P29466, P29594, P42573, P42574, P42575, P43527, P45436, P49662, P51878, P55210, P55211, P55212, P55213, P55214, P55215, P55865, P55866, P55867, P70343, P70677, P89116, P97864, Q14790, Q60431, Q92851, Q98943
C25	O33441, P28784, P95493, Q51816, Q51838, Q51839, Q51844

Sequences not in families C11, C13, C14 or C25
O22161, O28117, Q53839

Sequences not retrieved although recognised members of families C11, C13, C14 or C25

C11	None
C13	O24539, P09841
C14	P42576
C25	O07442, P72194, Q51817, Q51818

The two motifs were formulated according to the conventions of PROSITE [27] and used to search the SWISSPROT and TREMBL databases (10 August 1998; 223 599 sequences) by use of the SCANPROSITE program at the ExPASy server [27]. Motif 1 retrieved the 37 sequences underlined, all of which are recognised members of families C11, C13, C14 or C25. Motif 2 retrieved 66 sequences of which 63 are recognised members of these families, but three are not. Seven sequences of known members of families C11, C13, C14 or C25 were not retrieved by either motif. These were O24539, *Vicia narbonensis* legumain (with Cys-147); P09841, *Schistosoma mansoni* haemoglobinase (with Cys191Asn replacement); P42576, isoform of the positive P42575; O07442, P72194 and Q51817, Q51818, aberrant forms of gingipain. Five sequence fragments that did not overlap the consensus sequence were ignored.

3.4. Four families in clan CD

We believe that these findings represent strong evidence that the four families of cysteine peptidases C11, C13, C14 and C25 are evolutionarily related, albeit very distantly, and that they retain topological similarities in the catalytic sites. Testable predictions from this proposal are that the same histidine residue (His-150 in Fig. 3) will be found experimentally to be essential for catalysis in clostripain and gingipain R, and that when three-dimensional structures are determined for representatives of families C11, C13 and C25 they will show similarities to those of the caspases.

Family C14 has already been assigned to clan CD in the classification of the MEROPS database [9] on the grounds of its distinctive protein fold, and has been the only family in this clan, but we now propose that families C11, C13 and C25 should be added to clan CD. One of the purposes of assembling clans of peptidases is the expectation that the evolutionary and structural similarities between the peptidases in related families will tend to be reflected in similar enzymological properties and sometimes even functions. The enzymes of clan CD do indeed have striking similarities in catalytic activity. All are very specific endopeptidases, with a strict requirement for the side chain of the P1 amino acid residue that provides the carbonyl group in the scissile peptide bond: arginine for clostripain (family C11), asparagine for legumain (C13), aspartate for the caspases (C14) and arginine or lysine for gingipain R or gingipain K, respectively, in family C25. This type of specificity is not characteristic of the more familiar but quite unrelated cysteine peptidases that are homologues of papain in clan CA, which generally show broad specificity with greatest selectivity in subsite S2 [22,23]. There are also similarities in inhibition characteristics. All of the peptidases in the families now assigned to clan CD are resistant to irreversible inhibition by compound E-64 [L-3-carboxy-2,3-*trans*-epoxypropionyl-leucylamido-(4-guanidino)butane], which has often been thought of as a general reagent for cysteine endopeptidases. (Reversible inhibition of gingipain R and clostripain is attributable to the fact that E-64 is an arginine analogue [24,25].) However, the cysteine endopeptidases of clan CD are susceptible to peptide aldehyde and peptidylchloromethane inhibitors, and it now appears that it may be possible to take advantage of successes in the structure-based design of inhibitors for the caspases to design inhibitors for enzymes in the other families, modifying specificity simply by

matching the different P1 requirements in a similar active site geometry.

3.5. Evolution of clan CD

The families of clan CD are widely distributed. Clostripain, the only known member of family C11, occurs in Gram-positive bacteria, and the gingipains of family C25 in Gram-negative bacteria. In contrast, members of family C13 are known from plants, fungi and animals, and the caspases of family C14 are known only from animals. The idea that these four families are derived from a common ancestor raises the question of the route by which the caspases and the legumains reached mammals following the origin of the clan presumably in the prebiotic genote. The eukaryote sequences are very different from those in bacteria, and thus give no indication of horizontal transfer from bacteria to eukaryotes. None of the four families is yet reported to be represented in archaea or protista, but it seems that representatives of the clan must have existed there at some time, and we suggest they may yet be found in the modern members of these groups.

Acknowledgements: This work was supported by the Medical Research Council (UK). We thank Drs Pam Dando and Harald Thidemann Johansen for valuable comments.

References

[1] Kembhavi, A.A., Buttle, D.J., Knight, C.G. and Barrett, A.J. (1993) Arch. Biochem. Biophys. 303, 208–213.
[2] Hara-Nishimura, I., Kinoshita, T., Hiraiwa, N. and Nishimura, M. (1998) J. Plant Physiol. 152, 668–674.
[3] Dalton, J.P. and Brindley, P.J. (1998) in: Handbook of Proteolytic Enzymes (Barrett, A.J., Rawlings, N.D. and Woessner, J.F., Eds.) pp. 749–754, Academic Press, London.
[4] Min, W. and Jones, D.H. (1994) Nat. Struct. Biol. 1, 502–504.
[5] Chen, J.-M., Dando, P.M., Rawlings, N.D., Brown, M.A., Young, N.E., Stevens, R.A., Hewitt, E., Watts, C. and Barrett, A.J. (1997) J. Biol. Chem. 272, 8090–8098.
[6] Chen, J.-M., Dando, P.M., Stevens, R.A.E., Fortunato, M. and Barrett, A.J. (1998) Biochem. J. 335, 111–117.
[7] Manoury, B., Hewitt, E.W., Morrice, N., Dando, P.M., Barrett, A.J. and Watts, C. (1998) Nature, in press.
[8] Rawlings, N.D. and Barrett, A.J. (1998) Nucleic Acids Res., in press.
[9] Rawlings, N.D. and Barrett, A.J. *MEROPS* – the peptidase database. Release 3.03. http://www.bi.bbsrc.ac.uk/Merops/Merops.htm
[10] Picard, V., Ersdal-Badju, E., Lu, A. and Bock, S.C. (1994) Nucleic Acids Res. 22, 2587–2591.
[11] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[12] Götz, B. and Klinkert, M.-Q. (1993) Biochem. J. 290, 801–806.
[13] Yu, J., Nagarajan, S., Knez, J.J., Udenfriend, S., Chen, R. and Medof, M.E. (1997) Proc. Natl. Acad. Sci. USA 94, 12580–12585.
[14] Benghezal, M., Benachour, A., Rusconi, S., Aebi, M. and Conzelmann, A. (1996) EMBO J. 15, 6575–6583.
[15] Porter, W.H., Cunningham, L.W. and Mitchell, W.M. (1971) J. Biol. Chem. 246, 7675–7675.
[16] Potempa, J. and Travis, J. (1998) in: Handbook of Proteolytic Enzymes (Barrett, A.J., Rawlings, N.D. and Woessner, J.F., Eds.) pp. 762–768, Academic Press, London.
[17] Gilles, A.-M., De Wolf, A. and Keil, B. (1983) Eur. J. Biochem. 130, 473–479.
[18] Walker, N.P.C., Talanian, R.V., Brady, K.D., Dang, L.C., Bump, N.J., Ferenz, C.R., Franklin, S., Ghayur, T., Hackett, M.C., Hammill, L.D., Herzog, L., Hugunin, M., Houy, W., Mankovich, J.A., McGuiness, L., Orlewicz, E., Paskind, M., Pratt, C.A., Reis, P., Summani, A., Terranova, M., Welch, J.P., Xiong, L. and Möller, A. (1994) Cell 78, 343–352.
[19] Pavloff, N., Pemberton, P.A., Potempa, J., Chen, W.C.A., Pike, R.N., Prochazka, V., Kiefer, M.C., Travis, J. and Barr, P.J. (1997) J. Biol. Chem. 272, 1595–1600.
[20] Rost, B. (1996) Methods Enzymol. 266, 525–539.
[21] Witte, V., Wolf, N. and Dargatz, H. (1996) Curr. Microbiol. 33, 281–286.
[22] Brömme, D., Bonneau, P.R., Lachance, P. and Storer, A.C. (1994) J. Biol. Chem. 269, 30238–30242.
[23] Storer, A.C. and Ménard, R. (1996) Perspect. Drug Discov. Design 6, 33–46.
[24] Kembhavi, A.A., Buttle, D.J., Rauber, P. and Barrett, A.J. (1991) FEBS Lett. 283, 277–280.
[25] Rangarajan, M., Smith, S.J.M., U, S. and Curtis, M.A. (1997) Biochem. J. 323, 701–709.
[26] Kraulis, P.J. (1991) J. Appl. Crystallogr. 24, 946–950.
[27] Bairoch, A. and Bucher, P. (1994) Nucleic Acids Res. 22, 3583–3589.

Pfam 10.0 (Saint L uis)
Home | Analyze a sequence | Browse Pfam | Keyword search | Taxonomy search | Swisspfam | Help

Starting search. Estimated time: 14 seconds (assuming all Wulfpack nodes are running). Please wait...

Pfam HMM search results, glocal+local alignments merged (Pfam_ls+Pfam_fs)

[Go here for an explanation of the format of the results]

Model	Seq-from	Seq-to	HMM-from	HMM-to	Score	E-value	Alignment	Description
!! asp	10	485	1	414	584.4	**7.5e-173**	glocal	Eukaryotic aspartyl protease
!! SapB_2	294	328	1	35	66.1	**7.8e-17**	glocal	Saposin-like type B, region 2
!! SapB_1	357	395	1	39	69.2	**8.9e-18**	glocal	Saposin-like type B, region 1

Alignments of top-scoring domains:

Format for alignment of query to Seed: Plain text

```
asp: domain 1 of 1, from 10 to 485: score 584.4, E = 7.5e-173
                RF    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                   *->fariplkkgkSlreklsekgvlldflvkhkanpltksrlgaaasSkg
                       +r+ lkk      kl+ k+ l+ +   ++ +pl ++rlg+   S +
 AP|  query     10    TFRVGLKKL-----KLDSKNRLAARVESKQEKPLRAYRLGD---SGD 48

                RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                   savepllnyldDaeYygtIsIGTPpQkFtVvFDTGSSDLWVPDsSvyCts
                   ++v++l nyl Da+Yyg+I IGTPpQkFtVvFDTGSS+LWVP  S++C++
       query     49 ADVVVLKNYL-DAQYYGEIAIGTPPQKFTVVFDTGSSNLWVP--SSKCYF 95

                RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                   sySaqTniACkshgtFdPSkSSTYkslGttIffsIsYGdGSSasGflgqD
                   s       AC+ h ++ +S+SSTY ++G++   +I+YG+G    Gf++ D
       query     96 S------LACLLHPKYKSSRSSTYEKNGKA-A-AIHYGTGA-IAGFFSND 136

                RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                   TVtvGGisvtnQqFGlatkEPGsfFvtavfDGILGlGfpsisavggssaf
                    VtvG+++v+ Q+F +atkEPG++Fv a+fDGILGlGf++is+++
       query    137 AVTVGDLVVKDQEFIEATKEPGITFVVAKFDGILGLGFQEISVGK----- 181

                RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

```
                tytpVfdnlksQGLIdspaFSvYLNsddgsaqasgGeiiFGGvDpskYtG
                 + pV++n+ +QGLI++p+FS++LN+++ ++  +gGe++FGGvDp++++G
query    182 -AAPVWYNMLKQGLIKEPVFSFWLNRNA-DEE-EGGELVFGGVDPNHFKG 228

          RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                sltwvpVtssDdegdivsqgyWqitldsitvggsaCHttfcssGcqAIlD
                ++t+vpVt         ++gyWq+ ++++ +gg    t+fc+sGc AI+D
query    229 KHTYVPVT---------QKGYWQFDMGDVLIGGAP--TGFCESGCSAIAD 267

          RF xxxxxxxxxxxxxxxxxxxxxxxxx
                TGTsLlygPssivskiakavGAsese.....................
                +GTsLl+gP++i++ i+ a+GA+ + +++ ++  ++ +++  +   ++++
query    268 SGTSLLAGPTTIITMINHAIGAAGVVsqqcktvvdqygqtildlllsetq 317

          RF
                ..................................................
                +++  ++ +  + +++++ + + ++  ++++ + +++ ++   +  +
query    318 pkkicsqiglctfdgtrgvsmgiesvvdkenaklsngvgdaacsacemav 367

          RF                                  xxxxxxxxxxxxxxxxxxx
                ............................GeYvvdCdsisslpdvtFf
                   +++ +++ ++++  +  ++  ++ +++ Ge  vdC+++s++p+v+ +
query    368 vwiqsqlrqnmtqerilnyvnelcerlpspmGESAVDCAQLSTMPTVSLT 417

          RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                igGkkitVPpsayvlqnsqggssPndiClsGfqssddipppggplwILGD
                igGk + + p++yvl++++g   P  +C+sGf + +d++pp gplwILGD
query    418 IGGKVFDLAPEEYVLKVGEG---PVAQCISGFIA-LDVAPPRGPLWILGD 463

          RF xxxxxxxxxxxxxxxxxxxxx
                vFLRsyYvVFDrdNnrvGlApa<-*
                vF+++y++VFD++N +vG+A+a
query    464 VFMGKYHTVFDFGNEQVGFAEA    485
```

align query/10-485 to asp(ls) Seed

```
SapB_2: domain 1 of 1, from 294 to 328: score 66.1, E = 7.8e-17
                *->sdqCkkfVdqYgpliidlLvsgldPkevCsklglC<-*
                   s+qCk++VdqYg++i+dlL+s+++Pk++Cs++glC
query    294    SQQCKTVVDQYGQTILDLLLSETQPKKICSQIGLC    328
```

align query/294-328 to SapB_2(ls) Seed

```
SapB_1: domain 1 of 1, from 357 to 395: score 69.2, E = 8.9e-18
                *->dilCelCemvVkevenlLkdnkTqeeIlkaLeklCdlLP<-*
                   d+ C++Cem+V++++ +L++n Tqe Il+++++lC++LP
query    357    DAACSACEMAVVWIQSQLRQNMTQERILNYVNELCERLP    395
```

align query/357-395 to SapB_1(ls) Seed

NEW! Phyl gen mic analysis f query using RIO.

Pfam 10.0 (Saint Louis)
Home | Analyze a sequence | Browse Pfam | Keyword search | Taxonomy search | Swisspfam | Help

Starting search. Estimated time: 16 seconds (assuming all Wulfpack nodes are running). Please wait...

Pfam HMM search results, glocal+local alignments merged (Pfam_ls+Pfam_fs)

[Go here for an explanation of the format of the results]

Model	Seq-from	Seq-to	HMM-from	HMM-to	Score	E-value	Alignment	Description
!! asp	32	507	1	414	539.9	**1.8e-159**	glocal	Eukaryotic aspartyl protease
!! SapB_2	319	353	1	35	56.7	**5.3e-14**	glocal	Saposin-like type B, region 2
!! SapB_1	379	417	1	39	57.9	**2.3e-14**	glocal	Saposin-like type B, region 1

Alignments of top-scoring domains:

Format for alignment of query to Seed: Plain text

```
asp: domain 1 of 1, from 32 to 507: score 539.9, E = 1.8e-159
                RF    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                   *->fariplkkgkSlreklsekgvlldflvkhkanpltksrlgaaasSkg
                       +ri lkk k    ++s+++ ++ fl+ +  +  +k+ +  +++ +
      query     32    TIRIGLKKRK---LDRSNRLASQLFLKNRGSHWSPKHYFR-LND-EN 73

                RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                   savepllnyldDaeYygtIsIGTPpQkFtVvFDTGSSDLWVPDsSvyCts
                   ++ +pl nyl Da+Yyg+I+IGTPpQkFtV+FDTGSS+LW+P  S++C+
      query     74 ADMVPLKNYL-DAQYYGDITIGTPPQKFTVIFDTGSSNLWIP--STKCYL 120

                RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                   sySaqTniACkshgtFdPSkSSTYkslGttIffsIsYGdGSSasGflgqD
                   s      +AC  h+++ +S+SS+Y+++G++   sI+YG+G   sG+++ D
      query    121 S------VACYFHSKYKASQSSSYRKNGKP-A-SIRYGTGA-ISGYFSND 161

                RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                   TVtvGGisvtnQqFGlatkEPGsfFvtavfDGILGlGfpsisavggssaf
                    V vG+i+v+ Q+F +at+EPG++F  a+fDGILGlGf++is+++
      query    162 DVKVGDIVVKEQEFIEATSEPGITFLLAKFDGILGLGFKEISVGN----- 206

                RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

```
                tytpVfdnlksQGLIdspaFSvYLNsddgsaqasgGeiiFGGvDpskYtG
                  tpV++n++  GL+++p+FS++LN+++      +gGei+FGGvDp +++G
query    207 -STPVWYNMVEKGLVKEPIFSFWLNRNP-KDP-EGGEIVFGGVDPKHFKG 253

          RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                sltwvpVtssDdegdivsqgyWqitldsitvggsaCHttfcssGcqAIlD
                 +t vpVt          +gyWq+ ++++++ g+    t+ c+ Gc AI+D
query    254 EHTFVPVT---------HKGYWQFDMGDLQIAGKP--TGYCAKGCSAIAD 292

          RF xxxxxxxxxxxxxxxxxxxxxxxxx
                TGTsLlygPssivskiakavGAsese.......................
                +GTsLl+gPs++++ i+ a+GA+   +++ +   ++ +++  ++   +++
query    293 SGTSLLTGPSTVITMINHAIGAQGIVsreckavvdqygktmlnsllaqed 342

          RF
                ..................................................
                +++  ++ +    +++++ + + ++  ++++++  ++   +  +
query    343 pkkvcsqigvcaydgtqsvsmgiqsvvddgtsgllnqamcsacemaavwm 392

          RF                                xxxxxxxxxxxxxxxxxxxxx
                ............................GeYvvdCdsisslpdvtFfigG
                +++ ++++++++      +  ++ ++++++  vdC   ss+p vtF+igG
query    393 eseltqnqtqerilayaaelcdhiptqnQQSAVDCGRVSSMPIVTFSIGG 442

          RF xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
                kkitVPpsayvlqnsqggssPndiClsGfqssddipppggplwILGDvFL
                 ++ ++p +y+++ ++g +s   +C+sGf + +di+pp gplwILGD F+
query    443 RSFDLTPQDYIFKIGEGVES---QCTSGFTA-MDIAPPRGPLWILGDIFM 488

          RF xxxxxxxxxxxxxxxxxxx
                RsyYvVFDrdNnrvGlApa<-*
                + y++VFD++  rvG+A+a
query    489 GPYHTVFDYGKGRVGFAKA    507
```

align query/32-507 to asp(ls) Seed

```
SapB_2: domain 1 of 1, from 319 to 353: score 56.7, E = 5.3e-14
                *->sdqCkkfVdqYgpliidlLvsgldPkevCsklglC<-*
                   s +Ck++VdqYg ++++ L++++dPk+vCs++g+C
query    319    SRECKAVVDQYGKTMLNSLLAQEDPKKVCSQIGVC    353
```

align query/319-353 to SapB_2(ls) Seed

```
SapB_1: domain 1 of 1, from 379 to 417: score 57.9, E = 2.3e-14
                *->dilCelCemvVkevenlLkdnkTqeeIlkaLeklCdlLP<-*
                   +++C++Cem+ +++e  L +n+Tqe Il++  +lCd++P
query    379    QAMCSACEMAAVWMESELTQNQTQERILAYAAELCDHIP    417
```

align query/379-417 to SapB_1(ls) Seed

NEW! Phyl gen mic analysis of query using RIO.

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 5,283,184
DATED : February 1, 1994
INVENTOR(S) : Jorgensen; Richard A.; Napoli, Carolyn A.

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Title page,
After "The portion of the term of this patent subsequent to", "July 27, 2010" should be replaced with -- March 30, 2009 --.

Signed and Sealed this

Twenty-eighth Day of August, 2001

Attest:

Nicholas P. Godici

Attesting Officer

NICHOLAS P. GODICI
Acting Director of the United States Patent and Trademark Office

US005283184A

United States Patent [19]

Jorgensen et al.

[11] Patent Number: 5,283,184

[45] Date of Patent: * Feb. 1, 1994

[54] GENETIC ENGINEERING OF NOVEL PLANT PHENOTYPES

[75] Inventors: Richard A. Jorgensen; Carolyn A. Napoli, both of Davis, Calif.

[73] Assignee: DNA Plant Technology Corporation, Mt. Kisco, N.Y.

[*] Notice: The portion of the term of this patent subsequent to Jul. 27, 2010 has been disclaimed.

[21] Appl. No.: 687,550

[22] Filed: Apr. 17, 1991

Related U.S. Application Data

[63] Continuation-in-part of Ser. No. 501,076, Mar. 29, 1990, and a continuation-in-part of Ser. No. 331,338, Mar. 30, 1989, Pat. No. 5,034,323.

[51] Int. Cl.[5] C12N 15/29; C12N 15/82; A01H 4/00

[52] U.S. Cl. 435/172.3; 800/205; 800/DIG. 67; 800/DIG. 12; 435/320.1; 935/30; 935/35; 935/64; 935/67

[58] Field of Search ..... PLT/82.1, 77, 80, 82.4; 800/200, 205, DIG. 67, DIG. 12; 435/172.3, 320.1; 935/35, 64, 67, 30

[56] References Cited

U.S. PATENT DOCUMENTS

5,034,323 7/1991 Jorgensen et al. 435/172.3

FOREIGN PATENT DOCUMENTS

0218571 4/1987 European Pat. Off. .
0257993 3/1988 European Pat. Off. .

OTHER PUBLICATIONS

Niesbach-Klosgen et al. (1987) J. Mol. Evol. 26:213-225.
van der Krol et al. (Jun. 1988) Nature 333:866-869.
van der Krol et al. (1990) Plant Molecular Biology 14:457-466.
Meyer, et al. (1987) Nature 330: 677-678.
Fillati, et al. (Jul. 1987) Bio/Technology 5:726-730.
Martin; et al. (1985) The EMBO Journal 4(7):1625-1630.
Reif, et al. (1985) Mol. Gen. Genet. 199:208-215.
Reeves et al., (1986) Gen. and Cytogen., Biol. Abstr. 83(2): AB-454-455, 13846 (1986).
Krishnan et al. (1986) Plant Phys. Biochem. and Biophy., Bi l. Abstr. 82(7):AB-855, 67053.
Hansen (1987) Carlsberg Res. Commun., 52:381-392.
Scherer et al. (1987) Plant Mol. Biol. 9:127-134.
Horsch et al. (1985) Science, 227:1229-1231.
Hall et al. (1987) New Phytologist, 103:33-43.
Mol et al. (1983) Mol. Gen. Genet. 192:424-429.
Forkmann et al. (1979) Planta, 144:189-192.
Dunsmuir et al. (1987) Plant Biotechnology (T. Mabry, Ed.) IC2 Institute, 107-119.
Martin et al. (1985) EMBO J., 4:1625-1630.
Van Tunen et al. (1988) EMBO J., 7:1257-1263.
Spribille et al. (1982) Planta, 155:176-182.
Koes et al. (1987) Plant Mol. Biol., 10:159-169.
Koes et al. (1986) Nucleic Acids Res., 14:5229-5239.
Reimold et al. (1983) EMBO J., 2:1801-1805.
Kreuzaler et al. (1983) Proc. Natl. Acad. Sci. USA, 80:2591-93.
Fillatti et al. (1987) Bio/Technology 5:726-730.
Wiering et al. (1984) "Genetics of Flower and Pollen Colors" in Petunia, Monographs on Theoretical and Applied Genetics, vol. 9 (K. C. Sink, Ed.), Springer Verlag, Berlin.
Martin et al. (1987) "The Control of Floral Pigmentation in Antirrhinum Majus" in Developmental Mutants in Higher Plants, (Thomas & Grierson, Eds.) Cambridge University Press.
Harborne (1976) "Functions of Flavonoids in Plants" in Chem. and Biochem. of Plant Pigments (T. Goodwin, Ed.) 2nd Ed., vol. 1, Academic Press, London, p. 737.
Hahlbrock (1981) "Flavornoids" in the Biochem. of Plants (E. Conn, Ed.) vol. 7, Academic Press.
Ozeki et al. (1987) Physiologia Plantarum, vol. 69, 128.
Promega 1988/89 Catalog (Wisconsin, U.S.A., page 2).

Primary Examiner—Che S. Chereskin
Attorney, Agent, or Firm—Townsend and Townsend Khourie and Crew

[57] ABSTRACT

Methods are provided for producing plants exhibiting one or more desired phenotypic traits. In particular, transgenotes are selected that comprise a DNA segment operably linked to a promoter, wherein transcription products of the segment are substantially homologous to corresponding transcripts of endogenous flavonoid biosynthetic pathway genes.

31 Claims, 4 Drawing Sheets

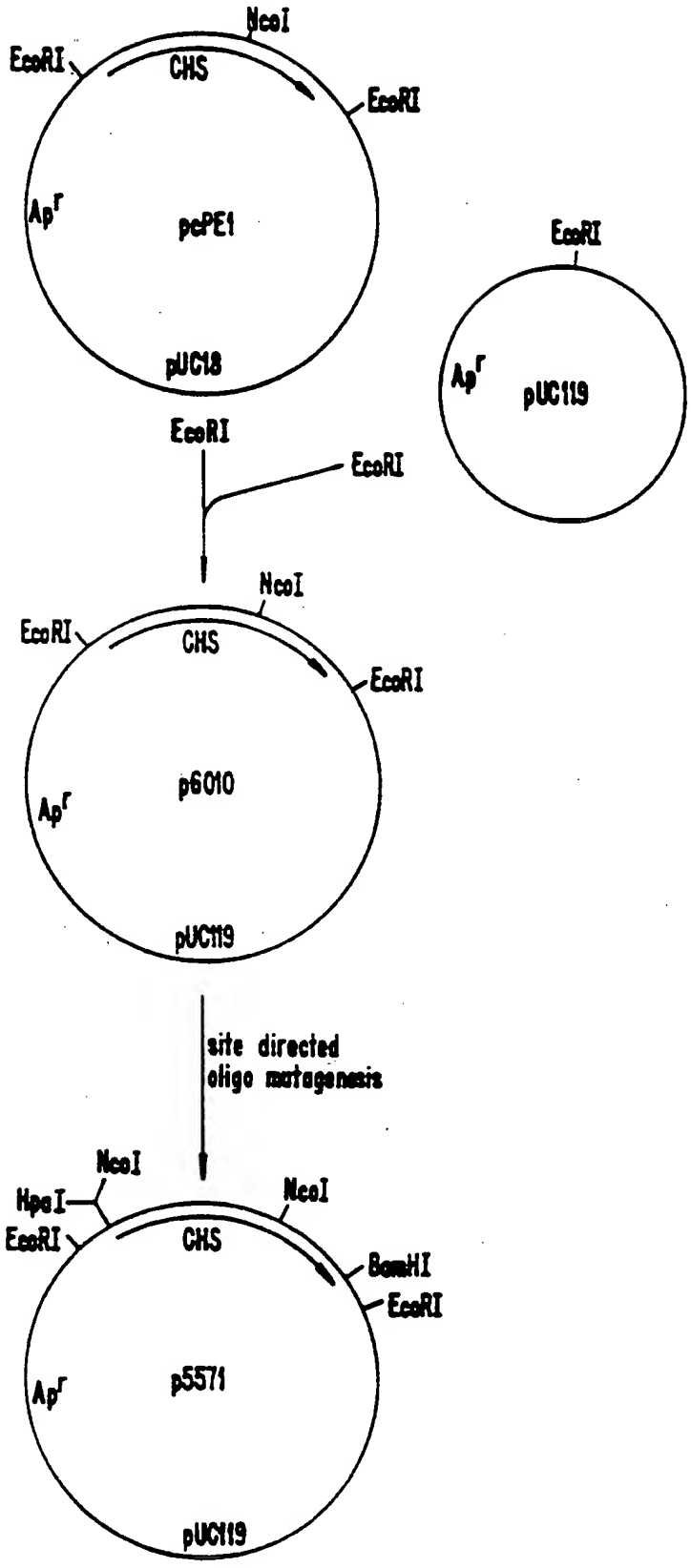

NcoI
EcoRI
CHS
EcoRI
Ap^r
pePE1
pUC18
EcoRI
Ap^r
pUC119
EcoRI
EcoRI
NcoI
EcoRI
CHS
EcoRI
p6010
Ap^r
pUC119
site directed
oligo mutagenesis
NcoI
HpaI
NcoI
EcoRI
CHS
BamHI
EcoRI
Ap^r
p5571
pUC119

FIG._1A.

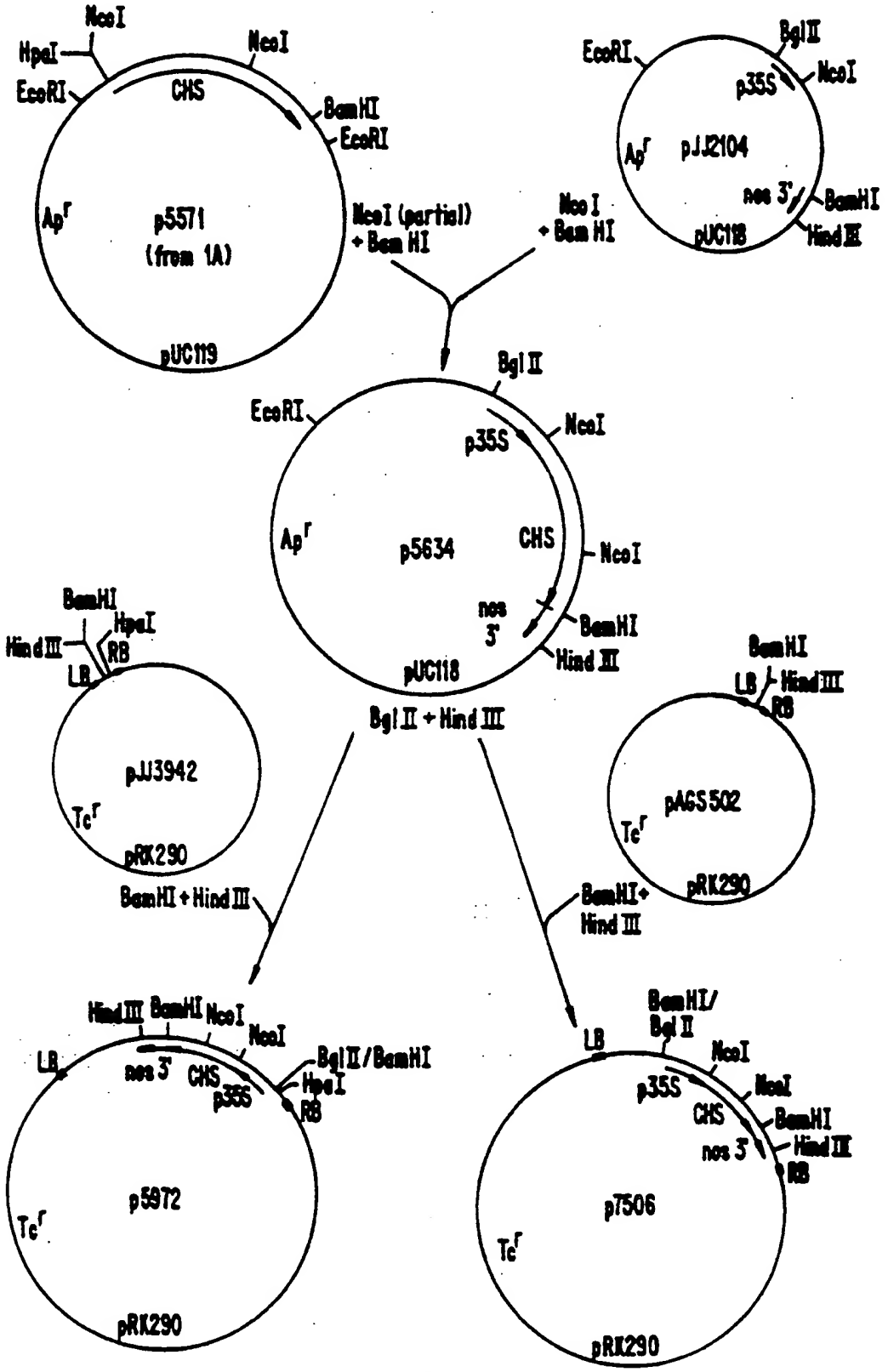

NcoI
HpaI
EcoRI
NcoI
CHS
BamHI
EcoRI
Ap^r
p5571
(from 1A)
pUC119
NcoI (partial)
+ Bam HI
EcoRI
BglII
p35S
NcoI
Ap^r
pJJ2104
nos 3'
BamHI
pUC118
HindIII
NcoI
+ Bam HI
BglII
EcoRI
NcoI
p35S
Ap^r
p5634
CHS
NcoI
nos
3'
BamHI
pUC118
HindIII
BglII + HindIII
BamHI
HpaI
HindIII
RB
LB
pJJ3942
Tc^r
pRK290
BamHI + HindIII
BamHI
HindIII
LB
RB
pAGS502
Tc^r
pRK290
BamHI+
HindIII
HindIII
BamHI
NcoI
NcoI
LB
nos 3'
CHS
p35S
BglII/BamHI
HpaI
RB
p5972
Tc^r
pRK290
BamHI/
BglII
LB
NcoI
p35S
NcoI
CHS
BamHI
nos 3'
HindIII
RB
p7506
Tc^r
pRK290

FIG._1B.

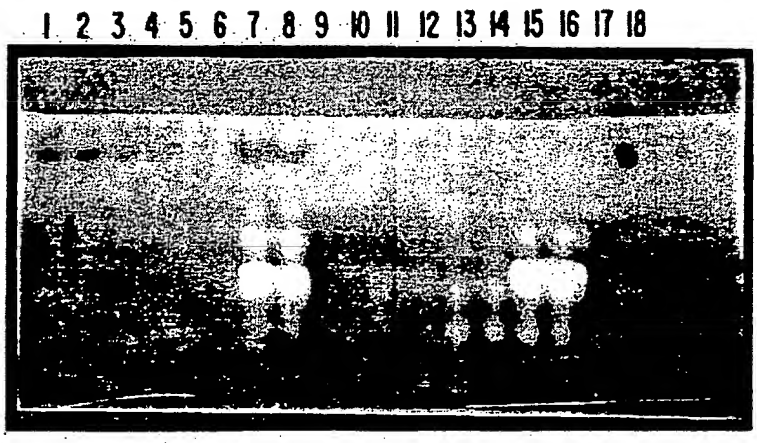


FIG. 2.

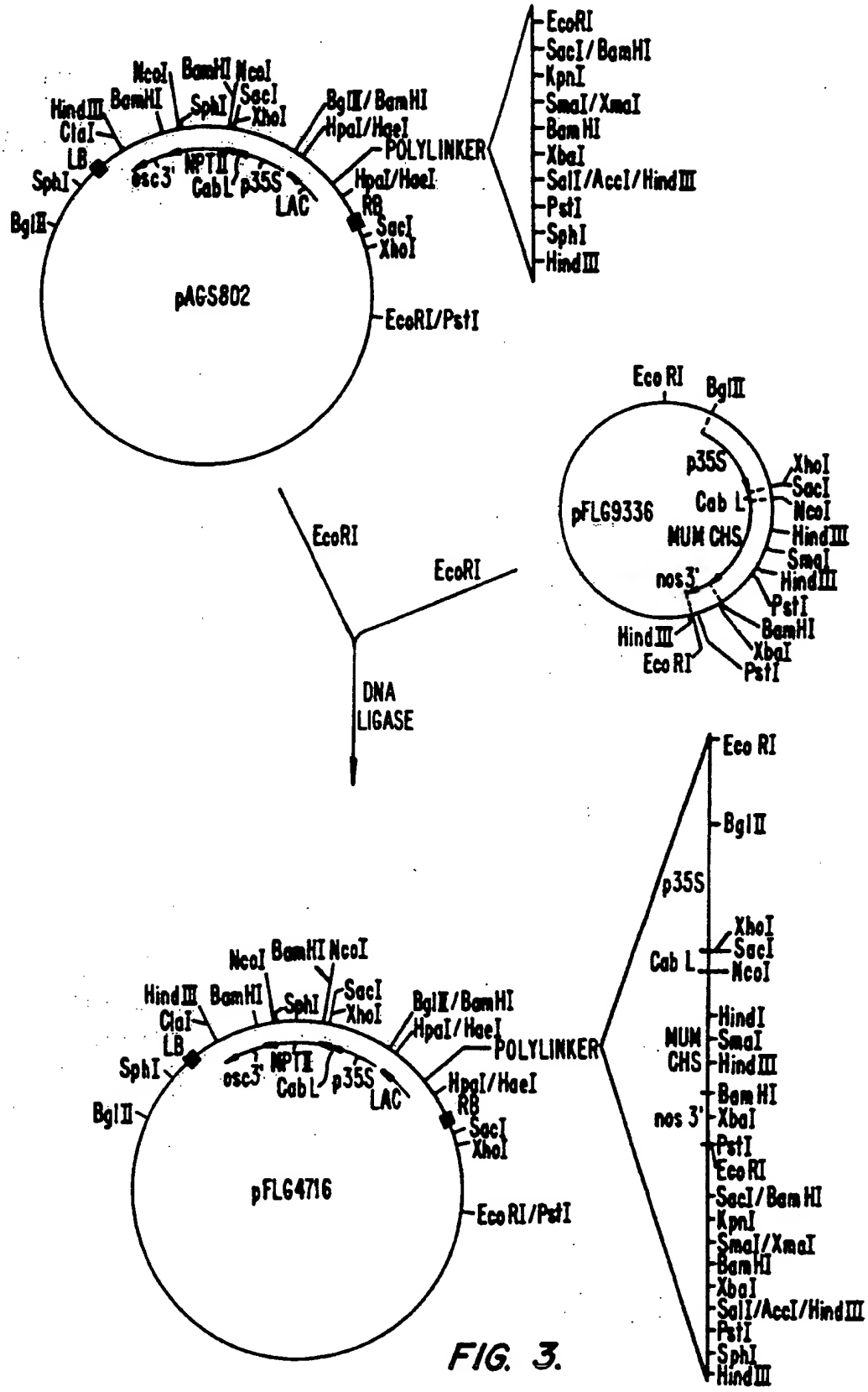

pAGS802
NcoI
BamHI
NcoI
SacI
XhoI
SphI
HindIII
BamHI
ClaI
LB
SphI
BglII
osc3'
NPTII
CabL
p35S
LAC
BglII/BamHI
HpaI/HaeI
POLYLINKER
HpaI/HaeI
RB
SacI
XhoI
EcoRI/PstI
EcoRI
SacI/BamHI
KpnI
SmaI/XmaI
BamHI
XbaI
SalI/AccI/HindIII
PstI
SphI
HindIII
EcoRI
EcoRI
DNA
LIGASE
pFLG9336
EcoRI
BglII
p35S
CabL
MUM CHS
nos3'
XhoI
SacI
NcoI
HindIII
SmaI
HindIII
PstI
BamHI
XbaI
PstI
HindIII
EcoRI
pFLG4716
NcoI
BamHI
NcoI
HindIII
BamHI
SphI
SacI
XhoI
ClaI
LB
SphI
BglII
osc3'
NPTII
CabL
p35S
LAC
BglII/BamHI
HpaI/HaeI
POLYLINKER
HpaI/HaeI
RB
SacI
XhoI
EcoRI/PstI
EcoRI
BglII
p35S
CabL
MUM CHS
nos 3'
XhoI
SacI
NcoI
HindI
SmaI
HindIII
BamHI
XbaI
PstI
EcoRI
SacI/BamHI
KpnI
SmaI/XmaI
BamHI
XbaI
SalI/AccI/HindIII
PstI
SphI
HindIII

FIG. 3.

GENETIC ENGINEERING OF NOVEL PLANT PHENOTYPES

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a Continuation-in-part of U.S. Ser. No. 07/501,076, filed Mar. 29, 1990, and claims priority from PCT application no. US89/03565 filed Aug. 18, 1989 and U.S. Ser. No. 07/331,338 filed Mar. 30, 1989, now U.S. Pat. No. 5,034,323.

FIELD OF THE INVENTION

This invention relates generally to the use of recombinant DNA methods for genetically altering plants, and more particularly, to improved means for altering plant phenotypes, such as color patterns and color intensity of flowers and other plant parts.

BACKGROUND OF THE INVENTION

Controlling metabolic pathways in plants has long been a goal of horticulturists. See, e.g., Bonner and Varner (1976) *Plant Biochemistry,* Academic Press, New York, which is incorporated herein by reference. The advent of recombinant DNA technology has provided new approaches to reaching that goal. While significant progress has been made in understanding gene regulation, control of plant gene regulation remains at a relatively early state of development.

The water soluble pigment flavonoids are significant in their contribution to the coloration and other properties of higher plants. For example, the flavonoids are responsible for most orange, scarlet, crimson, mauve, violet and blue colors, and contribute significantly to yellow, ivory and cream colored flowers. See, Harborne, (1976) *Chemistry and Biochemistry of Plant Pigments,* 2d ed., Goodwin (Ed.) Acad. Press, London. The most important of the pigment molecules are the anthocyanins, particularly pelargonidin, cyanidin and delphinidin. These are the darker colored pigments responsible for the orange-red, magenta and mauve colors, respectively. The other major flavonoid types, the chalcones, isomeric flavanones, flavones and flavonols are light colored and tend to have relatively smaller effects on intensity or patterns of color.

The functions of these pigments extend well beyond coloration of flowers, however. The pigments also color fruits, leaves and other plant parts, and importantly provide plants with UV protection, as well as protection against herbivores and microbes. Other uses include allelopathy and even some pharmaceutical applications.

The biosynthetic pathways of these various pigments have been extensively studied in many different plant species. The chalcones and aurones are products requiring only the initial biosynthetic enzymes, being direct products of the earliest precursors. The flavones and flavonols are intermediate, and the anthocyanins are products requiring substantial modifications from the initial precursors. All of these products are dependent upon the activity of the initial enzyme chalcone synthase (CHS), which catalyses the production of chalcone from three molecules of malonyl-Coenzyme A and one molecule of coumaroyl-Coenzyme A.

Essentially, all of these phenotypic traits have naturally evolved coordinately with constraints related to plant reproduction. For example, the appearance of a flower has generally resulted from the requirement to attract insects who assist in the pollination process essential for the sexual reproduction f the higher plants. Of course, the decorative and ornamental features impart to flowers a significant commercial value.

Mankind has traditi nally intervened in some f the natural processes by, e.g., simply selecting particular flower colors and patterns which might otherwise not have survived in nature. Breeders routinely generate new and unusual flower phenotypes by a variety of time-tested breeding methods. The classical techniques for breeding improved plants, such as different flower varieties with altered flower color intensities or color patterns, typically required natural genetic variability within the experimental gene pool of the species and its relatives. More recently, the generation of variability by induction of mutations has been utilized. Breeders then select among the resulting population those products exhibiting interesting phenotypes, for further characterization.

Unfortunately, the induction of mutations to generate diversity often involves chemical mutagenesis, radiation mutagenesis, tissue culture techniques, or mutagenic genetic stocks. These methods provide means for increasing genetic variability in the desired genes, but frequently produce deleterious mutations in many other genes. These other traits may be removed, in some instances, by further genetic manipulation (e.g., backcrossing), but such work is generally both expensive and time consuming. For example, in the flower business, the properties of stem strength and length, disease resistance and maintaining quality are important, but often initially compromised in the mutagenesis process.

As noted, the advent of recombinant DNA technology has provided horticulturists with additional means of modifying plant genomes. While certainly practical in some areas, to date genetic engineering methods have had limited success in modifying the flavonoid biosynthetic or other pathways. Recently, the inhibition of flower pigmentation with a constitutively expressed "anti-sense" chalcone synthase gene has been reported (Van der Krol et al., (1988) Nature 333:866–869).

Thus, there exists a need for improved methods for producing plants with desired phenotypic traits. In particular, these methods should provide general means for phenotypic modification, and may lessen or eliminate entirely the necessity for performing expensive and time-consuming backcrossing.

SUMMARY OF THE INVENTION

In accordance with the present invention, methods and compositions are provided for producing plants exhibiting one or more desired phenotypic or genotypic traits. The invention is based in part on the surprising discovery that plants exhibiting desired trait(s) can be selected from transgenotes comprising a nucleic acid segment operably linked to a promoter, wherein transcription products of the segment are substantially homologous to corresponding transcripts of an endogenous gene. The transgenotes are grown into plants, such as flowering plants capable of exhibiting novel traits, including a reduction in color intensity, an altered pattern color, or a change in basic color of the plant flowers or other plant organs.

The invention further embraces the introduction f ne or more metabolic enzyme gene regions, e.g., flavonoid biosynthetic pathway gene regions, under the control of promoter regions (either native or heterologous),

int dicots and other plants in which the gene is endogen us. In particular, the invention comprises plants, such as those f the genera Petunia and Chrysanthemum, wherein the plant is gr wn from a cell transformed with a sequence which is transcribed int an RNA sequence substantially homolog us (ther than anti-sense) t a desired gene. In a preferred embodiment the gene is a flavonoid biosynthetic pathway gene. DNA or RNA equivalents are introduced into plants in a way to produce more of the endogenous (already present) transcript, but not to produce solely an anti-sense transcript. This is preferably accomplished by using a DNA segment (natural or constructed) in which the promoter is positioned in the normal orientation at the 5' end of the encoding region so that a "sense" transcript (rather than antisense transcript) will be produced. The plant cells can be transformed with a variety of vectors, such as viral vectors, episomal vectors, shuttle vectors, Ti plasmid vectors and the like, all in accordance with well known procedures.

The invention also embraces methods for reducing expression of endogenous nucleic acid sequences coding for proteins acting in various biosynthetic or other enzyme pathways, such as the flavonoid biosynthetic pathway of a plant, the method comprising the step of introducing into a cell of the plant a DNA sequence substantially homologous to the endogenous sequence and under the operational control of a promoter sequence, such as a cauliflower mosaic virus sequence. The DNA segment will be sufficient to introduce the repression effect into the plant and typically comprises at least about 50 nucleotides but may be a full length gene. In addition, the invention comprises the methods of preparing and using the various DNA constructs of the present invention.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1A shows the constructions of p5571.

FIG. 1B shows the construction of p5972 and p7506.

FIG. 2. Flavonoid extracts of petunia corollas chromatographed on crystalline cellulose chromatography places in Forestal solvent (acetic acid: conc. HCL: water, 30:3:10). Lanes: 1,2- 21838; 3,4- 21841; 5,6- W80; 7,8-V26; 9,10- W37; 11,12- W85; 13,14- R77; 15,16-R27; 17-caffeic acid; 18- coumaric acid. The plate was photographed using UV light from below, and photographed with a red filter. Anthocyanins and coumaric acid can be detected readily this way, but caffeic acid and flavonols are not detected well.

FIG. 3 shows the construction of plasmids useful in modifying chrysanthemum flower pigments.

DETAILED DESCRIPTION

The present invention provides novel methods for producing plants, and embraces the plants so produced, and methods of their use. The invention is based in part on the discovery that a reduction in expression (i.e., repression) of a cellular gene product may be attained upon introduction into the cell of a nucleic acid fragment, e.g., a flavonoid biosynthetic pathway gene sequence, that is ultimately transcribed to yield a mRNA transcript substantially homologous to a portion of the gene's transcript. The introduced transcript is preferably produced prior to native transcript (if any), but may be produced simultaneously with native transcript production. Depending on the time and amount of transcript produced in a transgenote, a plant grown from it will exhibit a variety of different phenotypic traits. In particular, selecting plants with varying phenotypes, for instance col r patterns and intensity, typically without harming ther desirable plant characteristics, can be readily achieved in accordance with the present invention.

By way f example, and n t limitation, an exemplary preferred embodiment of the present inventi n entails introducing a full-length chalcone synthase (CHS) coding sequence in an orientation which would be operably linked to a cauliflower mosaic virus promoter into *Petunia hybrida* cells. These transgenotes are grown into plants and variations in flower coloration are selected. The modified flowers exhibit substantially all of the characteristics of the native *Petunia hybrida* plants. In another aspect of the present invention, those skilled in the art will readily appreciate that additional traits (i.e., other than proteins in the flavonoid biosynthetic pathway), additional plant nucleic acid sequences and the like may be readily substituted in accordance with the following guidelines.

Traits

A variety of traits are selectable with appropriate procedures and sufficient numbers of transgenotes. Such traits include, but are not limited to, visible traits, environmental or stress related traits, disease related traits, and ripening traits. The repressive effect is applicable to a variety of genes expressed in plants including, for example, genes responsible for the synthesis or metabolism of peptides, proteins, fatty acids, lipids, waxes, oils, starches, sugars, carbohydrates, flavors, odors, fragrances, toxins, carotenoid pigments, hormones, cell wall polymers, gene regulatory molecules, flavonoids, storage proteins, phenolic acids, coumarins, alkaloids, quinones, lignins, glucosinolates, tannins, aliphatic amines, celluloses, polysaccharides, glycoproteins and glycolipids.

For instance, an alteration in the production of fatty acids or lipids can be engineered (and fatty acid composition of, e.g., an oil-producing plant thus altered) by blocking synthesis of a specific chain elongation or desaturation enzyme. Also, the synthesis of starch can be reduced and sugars accumulated (and sugar content of, e.g., an edible plant thus altered) by blocking an enzyme required for starch synthesis. Similarly, fragrant molecules can be released from cells (thus altering the scent characteristics of, e.g., ornamental flowers) by blocking the enzymes responsible for glycosylation of such molecules.

Among the easiest to select are the flavonoid genes, giving rise to visible traits. In particular, the traits of color intensity, color hue and color pattern are subject to the repression effect.

In addition to variation for color amongst flowers, there is also variation amongst white colored flowers for the block in the flavonoid biosynthetic pathway. For example, in petunia there are mutations which block the pathway at the dihydroflavonol synthase step, so that dihydroflavonols and flavonols accumulate in the flowers (Beld et al., 1989, Plant Mo. Biol., 13, 491-502, which is incorporated herein by reference). In snapdragon, there are mutations which block the pathway at the chalcone synthase step, so that precursors to these reaction accumulate (for example, caffeic or coumaric acids). In both snapdragon and petunia, mutations at the flavanone-3-hydroxylase step lead to the accumulation of naringenin, resulting in white fl wers (e.g., Froemel et al., 1985, Theor. Appl. Gent., 70, 561). By

such analysis f the accumulated intermediates in white fl wers, it is possible to assess the point at which flavonoid biosynthesis has been blocked.

The class of genes within the flavon id biosynthetic pathway includes those nucleic acid sequences directly involved in reactions or control of reactions which synthesize or modify a flavonoid compound. Flavonoids are a class of compounds, numbering about 3000 whose functions in plants include coloration in flowers, fruits, leaves, and other organs. Examples of flavonoid biosynthetic genes include those for chalcone synthases, chalcone isomerases (CHI), flavanone 3-hydroxylases, dihydroflavonol reductases, flavanone 2-hydroxylases, dihydroflavanol 2-hydroxylases, flavonoid 3'-hydroxylases, flavonoid 5'-hydroxylases, flavonoid glycosyltransferases (including glucosyl transferases such as UDPG: flavonoid 3-O-glucosyl transferase and UDPG: flavonol 7-O-glucosyl transferase, and rhamnosyl transferases), flavonoid methyltransferases (such as SAM:anthocyanidin 3-(p-coumaroyl)-rutinoside-5-glucoside 3',5'-O-methyltransferase) and flavonoid acyltransferases. See, Hahlbrock (1981) *Biochemistry of Plants*, Vol. 7, Conn (Ed.); Weiring and de Vlaming (1984) "Petunia", K. C. Sink (Ed.), Springer-Verlag, New York. By way of example of the components of these pathways, phenylalanine is converted to cinnamic acid; cinnamic acid is converted to caffeic acid (in a side branch) and to coumaryl-coenzyme A; coumaryl-coenzyme A is converted with chalcone synthase to a tetrahydroxy chalcone; the tetrahydroxy chalcone is converted with chalcone isomerase to naringenin; and naringenin is converted in a subsequent series of steps to anthocyanins. Blockage or inactivation at a given stage leads to build-up of precursors and side chain products of precursors. For instance, blockade at the chalcone synthase stage leads to build-up of coumaric acid and caffeic acid.

Anthocyanin pigmented flowers have colors throughout the range orange to red to purple to blue. Chalcones and aurones are yellow or orange. Flavones and flavonols are very light yellow, or "cream" colored. Flavanones are colorless. Elimination of anthocyanins and diversion of the pathway to flavone or flavonol production would create cream colored flowers. Shifts from blue to purple or purple to red or red to orange can be engineered by interfering with 3' or 5' hydroxylases of 2-hydroxylases. Interference with 2-hydroxylases can also reduce color intensity of anthocyanin pigmented plants. Interference with CHS would create white flowers and with chalcone isomerase would create yellow flowers. A wide variety of bi-color patterns can be created, the two colors being the color of the target plant before engineering and the color resulting from the expression of the introduced flavonoid gene. Types of patterns include: radial star-like patterns; picotee (white outer edge); white annular center; concentric colored rings; erratic, irregular patterns, e.g., variegated or blotchy. There are many variations on these patterns, some more attractive than others, some with sharp boundaries between colors, some with diffuse boundaries, some with linear boundaries, some with wavy, curved boundaries. Also lighter, solid colors are observed.

The present invention is also applicable to generate plants exhibiting modified absolute levels or relative proportions of vari us oil products. In particular, genes encoding enzymes involved in lipid metabolism may be used, in the methods described, to modify metabolism of ils and other lipids, such as fats. The ils are lipids which typically are liquid at room temperatures and typically will have various unsaturations in the fatty acid components and shorter lipid chains than the fats. See generally, Lehninger (1978) *Biochemistry* (2d Ed), W rth Publishers, New York; and Bonner and Varner (1976) *Plant Biochemistry* (3d Ed), Academic Press, New York. Much is also known about the applicable biosynthetic pathways at the genetic level. See, e.g., the following, and references cited therein: Shure et al. (198), *Cell* 35:225–233; Preiss et al., *Tailoring Genes for Crop Improvement* (Bruening et al., eds.), Plenum Press (1987), 133–152; Gupta et al. (1988), *Plant Mol. Biol.* 10:215–224; Olive et al. (1989), *Plant Mol. Biol.* 12:525–538; and Bhattacharyya et al. (1990), *Cell* 60:155–122, all of which are incorporated herein by reference.

Genes encoding various enzymes found in the pathways for production of oils are introduced into a target plant to alter the expression of important endogenous genes. The relative proportions or absolute content of various different types of oils may be thereby altered, and using information known in the art about sugar-starch biosynthetic pathways at the genetic level (see, e.g., citations on same above). Various enzymes of particular interest would include, among others, stearyl desaturase, acetyl transacylase, malonyl transacylase, β-ketoacyl ACP-synthetase, β-keto ACP-reductase, enoyl ACP-hydrase, acyl-ACP thioesterases and enoyl ACP-reductase.

The present invention is also applicable to the alteration of sugar metabolism or carbohydrate metabolism in plants. For example, genes encoding enzymes used in carbohydrate metabolism, e.g, in the metabolism of amylose, pectins, cellulose and cell walls, are used to regulate enzymatic expression and activity to modify sizes of pools of metabolic intermediates or kinetics of conversion, thereby changing the starch or sugar contents of various plants.

Genes encoding various enzymes in carbohydrate metabolism may be introduced into a target plant to alter the expression of various enzymes in the pathways. Various enzymes of particular interest in carbohydrate metabolism include phosphorylase, starch synthetase, Q-enzyme, sucrose-6-phosphate synthetase, sucrose-6-phosphate phosphatase, ADP-glucose pyrophosphorylase and various amylases. See generally, Bonner and Varner (Eds) (1976) *Plant Biochemistry* (3d ed.), Academic Press, New York. Much is also known about the applicable biosynthetic pathways at the genetic level. See, e.g., the following, and references cited therein: Stefansson et al. (1961), *Can. J. Plant Sci.* 41:218–219; Knowles et al. (1972), *Oil Crops of the World* (Robbelen et al., eds.), McGraw-Hill, 260–282; Hammond et al. (1983), *Crop Sci.* 23:192–197; Widstrom et al. (1984), *Crop Sci.* 24:1113–1115; Green et al. (1984), *Euphytica* 33:321–328; Graef et al. (1985), *Crop Sci.* 25:1076–1079; Somerville et al., *Recent Advances in Phytochemistry* (Conn, ed.), Plenum Press (1988), 19–44; Kunst et al. (1988), *PNAS U.S.A.* 85:4143–4147; and, Browse et al., "Strategies for Modifying Plant Lipid Composition" (1989), *Plant Gene Transfer* (Lamb, C. et al., eds.), Alan R. Liss, all of which are incorporated herein by reference.

Suitable sources for gene sequences usable in accordance with the present invention are plants, in particular higher plants. For example, virtually all f higher plants normally possess a flavonoid biosynthetic pathway f some type. In particular, any flavonoid pathway

which generates naringenin chalcone or compounds generated from naringenin chalcone which itself is generated fr m coumar yl-Coenzyme A and malonyl-Coenzyme A by chalcone synthase will be appropriate.

Introduced Nucleic Acid Sequences

The properties of the nucleic acid sequences are varied, and the preferred embodiments will describe a number of features which the person of skill in the art may recognize as not being absolutely essential, but clearly advantageous. These include isolation methods of the particular sequence to be introduced, certain features of the sequence and certain features of the associated vector, if any. Transcriptional expression of the introduced gene is generally important, and—without intending to be limited to a particular mechanism—additional production of a transcript relative to the normal expression of the endogenous form of the sequence is likely part of the underlying mechanism, especially prior to attainment of peak levels of endogenous gene expression. However, other mechanisms may be involved, for instance, mere physical presence of exogenous copies of these genes, e.g., in higher copy numbers or integrated into particular genetic locations, may contribute to the repressive effect.

RNA resulting from transcription shall be referred to herein on occasion as "transcript" or in some instances "mRNA". Typically, transcript which is processed (e.g., introns removed and 3' end polyadenylated) is referred to as mRNA ("messenger"). As used herein "homologous" means corresponding to (the same as). For example, RNA which is homologous to a gene, is RNA which corresponds to the coding strand sequence (with the normal exception of uracil for RNA in place of thymidine for DNA). Thus, cellularly produced "homologous mRNA", as used herein, is complementary to the template DNA strand of the gene.

mRNA may contain "coding regions" of codon triplets which encode for translation into amino acids making up the polypeptide. A primary transcript may contain both exons (typically containing the coding regions) and introns, the latter of which are often excised prior to translation.

Expression of an endogenous gene, e.g., a gene in the flavonoid pathway, yields varying levels of transcript depending on the type of cell and its developmental stage. During flower development, certain cells, e.g., cells that give rise to petal epidermal tissue, produce or begin to produce a transcript at a level which rises at or subsequent to flower meristem initiation. The transcript level reaches a peak later in flower development and eventually decreases. This rise and fall of transcript level may occur over a series of days, e.g., 7–14 days. The rise may also occur rapidly, e.g., over a period of hours, especially in the event of induction such as by UV or visible light (see, Mol et al., (1983) *Mol. Gen. Genet.* 192:424–429, which is incorporated herein by reference), particularly under natural summer light conditions (or the artificial equivalent thereof). For example, the transcript level is usually decreasing at the mature flower stage (flower maturation).

One proposed mechanism of repression would require that some transcription of the introduced sequence be produced, e.g., transcription of introduced DNA homologous to transcribed endogenous DNA (although transcription of introduced DNA homologous to untranscribed endogenous DNA may also be involved). Specifically, however, bserved genes' endogenous transcription levels dropped about 50-f ld in comparison to native conditions. The introduced genes' transcript levels were also l w with respect to the endogenous genes' levels, but the total was also lower. While the introduced sequence need not necessarily correspond t the final translated message or a part thereof, there are likely corresponding f rms of the mRNA which are functional in repression, but still contain parts of introns or only non-translated segments of the primary transcript of the normal endogenous sequence. Thus, the effect may occur where the introduced sequence contains no coding sequence per se, but only intron or untranslated sequences substantially homologous to sequences present in the primary transcript of the endogenous sequence. Also, while the present invention is not necessarily tied to any specific mechanism, the mechanism may involve methylation of both the introduced and the endogenous homologous sequences (methylation is known to be able to spread from a sequence to adjacent sequences).

The introduced sequence generally will be substantially homologous to the endogenous sequence intended to be repressed, such that the controlling elements recognize that the introduced sequence is present, the interaction results in the repressive effect. This minimal homology will typically be greater than about 65%, but a higher homology might exert a more effective repression of expression of the endogenous sequences. Substantially greater homology, or more than about 80% is preferred, though about 95% to absolute identity would be most preferred. Consequently, the effect should apply to any other proteins within a similar family of genes exhibiting homology or substantial homology. For example, the chalcone synthase protein may be encoded by one or more homologous genes which comprise the chalcone synthase gene family, and repression of one member of the family will typically serve to impose the same repressive effect on others of the family. Similarly, for example, chalcone synthase genes from other plant species may be utilized.

The introduced sequence, needing less than absolute homology, also need not be full length, relative to either the primary transcription product or fully processed mRNA. In this regard and in accordance with the present invention, transformation of partial or truncated genes also inhibits expression. Specifically, transformants of a tobacco plant with an 860 base pair fragment containing approximately the first two-thirds of a wild-type nopaline synthase gene under control of the CaMV 35S promoter showed suppression of the wild-type nos gene previously transformed into the tobacco Caltivar (see, Goring et al., *Proc. Natl. Acad. Sci. U.S.A.*, 88:1770–1774, at 1771 (1991), which is incorporated herein by reference). Similarly, tomato plants transformed with a 730 bp fragment (about 678 bp of coding sequence, representing about 51% of the total length; i.e., 1382 bp coding sequence plus 52 bp at 5' end) from a polygalacturonase (PG) cDNA (Grierson, et al., *Nucl. Acids. Res.* 14:8595–8603 (1986)) under the control of the CaMV 35S promoter showed inhibition of the endogenous PG gene during ripening (see, Smith, et al., *Mol. Gen. Genet.*, 224(3):477–481, at 478 (1990), which is incorporated herein by reference).

A higher homology in a shorter than full length sequence compensates for a longer less h mol g us sequence. Furthermore, the introduced sequence need not have the same intron r exon pattern, and h mology of non-coding segments will be equally effective. Nor-

mally, a sequence of greater than 50–100 nucleotides should be used, th ugh a sequence f greater than about 200–300 nucleotides w uld be preferred, and a sequence f greater than 500–1000 nucleotides w uld be especially preferred depending on the size of the endogenous gene.

It should be noted that since a full length coding sequence is unnecessary, it is possible to produce the same effect on multiple proteins using a single transformation by fusing multiple sequences together to coordinately repress various different genes. Assuming a sufficient number of introductions are made, the introduced sequence need not be linked to an operative promoter sequence. However, a promoter sequence would be preferred, particularly a partially or fully constitutive promoter. "Operably linked" refers to functional linkage between the affecting sequence (such as a promoter or 3' segments) and the controlled nucleic acid sequence. The same effect would be produced by the introduction of a promoter operably linked to the coding strand of an endogenous sequence. This can be effected by either the introduction of a promoter alone to a site operably linked to the target sequence, or by the reintroduction of a sequence of endogenous origin recombinantly attached to an operably linked promoter (resulting in a chimeric gene).

A heterologous sequence is one that originates from a foreign species, or, if from the same species, is substantially modified from its original form.

In considering the expected temporal stage of expression of the introduced gene, relevant factors include the type of promoter, the temporal pattern of the promoter, and the operation of the promoter in view of its position within the genome. A promoter which is expressed concurrently with or prior to the normal activation of the homologous endogenous sequence is preferred. A constitutive promoter is most preferred, such as the cauliflower mosaic virus promoter. This promoter is constitutive because its operation is relatively independent of the developmental stage of the cell in which it is contained. A regulated promoter, such as ones associated with the ribulose-1,5-bisphosphate carboxylase, the chlorophyll binding proteins or the glycine-rich root protein genes are also suitable. This control may be either temporal with respect to the developmental stage of the cell, or based upon differential expression by different parts or organs of the plant.

As referred to above, the operation of a promoter may vary depending on its location in the genome. Thus, a regulated promoter may operate differently from how it does in its normal location, e.g., it may become fully or partially constitutive.

It is preferred to have the DNA sequence linked to and situated at a distance from the promoter corresponding to the distance at which the promoter is normally most effective so as to ensure sufficient transcriptional activity. This distance should be within about 1000 nucleotides, preferably within about 500 nucleotides and more preferably within about 300 nucleotides of the translation initiation codon.

At the 3' end of the coding sequence, operably linked segments may also be included. Thus, it would be optimum to have a 3' untranslated region containing the polyadenylation site and any relevant transcription termination sites. A 3' sequence of less than about 1000 nucleotides is sufficient, about 500 preferred and about 300, or the length f the 3' untranslated tail of the endogenous sequence is more preferred.

If the introduced gene is an intact gene from the target plant r other plant species (meaning a complete gene containing coding sequences, intron, promoter, enhancers and other cis-acting regulat ry elements either upstream (5') or downstream (3') of the coding sequences), a fraction of independent transgenotes, depending on the gene, may carry the introduced gene in locations that result in abnormal expression, i.e., expression at abnormal times in development. If the introduced gene is a chimeric gene (meaning that one or more elements, such as a promoter, from another gene has been substituted for a component of the intact gene or added to the intact gene, including coding sequences fused to upstream and downstream sequences necessary or beneficial for expression) and is driven by a constitutive (fully or partially) promoter, then abnormal levels and times of expression will be achieved in a large fraction of transgenotes. If the introduced gene is a chimeric gene and is driven by a developmentally regulated promoter, depending on the promoter, some fraction of transgenotes will show abnormal levels and times of expression of the introduced gene. The strength of the promoter or other cis element can be the same, lower, or higher than the coding sequence's usual promoter. The timing in development can be earlier or the same.

While many of these improvements suggested are not essential, the efficiency of production of useful transgenotes may be significantly affected. Some of the transgenotes may be identical to the parental plants, others may have reduced amounts of colored or colorless flavonoids throughout the petals or other organs of interest. Others may have reduced amounts of flavonoids in certain cells or patches of cells or segments of petals or other organs resulting in regular or irregular patterns. Flowers on the same plant may even have different patterns. The likelihood of obtaining a desirable transgenote will depend upon the number of transgenotes screened and the efficiency of actual transformation and expression of the foreign nucleic acid sequence. Typically, at least about 25 to 50 transgenotes will be screened, but 100 to 500 or more may need to be screened before the described effect is seen.

The choice of nucleic acid to exert the described repressive effect is broad. Assuming appropriate selection procedures and sufficient numbers of transgenotes, a wide variety of plant genes could display this effect. For example, genes responsible for the synthesis or metabolism of peptides, proteins, fatty acids, lipids, waxes, oils, starches, sugars, carbohydrates, flavors, odors, fragrances, toxins, carotenoid pigments, hormones, cell wall polymers, gene regulatory molecules, flavonoids, seed storage proteins, phenolic acids, coumarins, alkaloids, quinones, lignins, glucosinolates, tannins, aliphatic amines, celluloses, polysaccharides, glycoproteins and glycolipids and particularly genes associated with plant pigmentation. Among the plant pigmentation genes are the flavonoid genes, and most particularly the chalcone synthase gene sequence.

These gene sequences may be isolated by standard procedures of hybridization of genomic or cDNA libraries by the methods described in Maniatis et al. (see below). Screening may be by (1) nucleic acid hybridization using homologous genes from ther rganisms, (2) probes synthetically produced to hybridize to particular sequences coding for known protein sequences, or (3) DNA sequencing and comparison to known sequences. Sequences for specific genes may be found, e.g., in GenBank, National Institutes of Health computer data-

base), or may be determined after isolation, typically using techniques as described.

Flavonoid genes, f r example, may be enriched in libraries by differential hybridization which requires that the mRNA of the target genes be expressed m re abundantly in one tissue than in another. Labelled RNA or cDNA from each tissue is hybridized to replicas of the library and tissue specific clones are identified and isolated. Screening can then be used to identify the target gene among the set of tissue specific genes (Kreuzaler et al., (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80:2591-2593).

Antibody screening of expression libraries with antibodies made against homologous proteins can select nucleic acid sequences which would code for homologous functions. Selection of sequences homologous to a known flavonoid biosynthetic pathway protein will enable isolation of other forms or equivalent forms from different sources.

Transposon tagging can assist in the isolation of the relevant gene. Transposon tagging typically involves a mutation of the target gene. A mutation is isolated in which a transposon has inserted into the target gene and altered the resulting phenotype. Using a probe for the transposon, the mutant gene can be isolated. Then, using the DNA adjacent to the transposon in the isolated mutant gene as a probe, the normal wild type allele of the target gene can be isolated (McLaughlin and Walbot (1987) *Genetics* 117:771-776; Dooner et al., (1985) *Mol. Gen. Genetics* 200:240-246; and Federoff et al., (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81:3825-3829).

However, as indicated above, the homology between the inserted gene and the endogenous gene need not be absolutely identical. Foreign homologous genes would also be subject to this same repression phenomenon. As stated, the repressive effect can occur with many different genes. It is exemplified herein, inter alia, with respect to flavonoid pathway genes.

Target Plants

As used herein, "plant" refers to either a whole plant, a plant part, a plant cell, or a group of plant cells. The class of plants which can be used in the method of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants. It includes plants of a variety of ploidy levels, including polyploid, diploid and haploid.

Although useful in regulating expression of many plant genes, the invention is easily characterized with particular application to plants which express the flavonoid pathway genes. At least some of the flavonoid pathway genes are essentially ubiquitous in higher plants; their products are found in flowers or other plant organs (such as leaves, stems, roots, tubers, bracts, sepals, fruits, vegetables) which are colored. These colors are provided largely by anthocyanin pigments, other flavonoid pigments, copigments, or colorless flavonoids synthesized from chalcone by the plant. See Hahlbrock, supra; Harborne, (1986) *Plant Flavenoids in Biology and Medicine: Biochemical Pharmacological and Structure Activity Relationships;* Harborne, (1976) *Chemistry and Biochemistry of Plant Pigments,* (2d ed.) Vol. 1, Goodwin (Ed.) Acad. Press.

Fruit (e.g., apples, cherries, plums, grapes), vegetable (e.g., eggplant, peppers, kale, lettuce, radishes, cauliflower) or ther edible plant part (e.g., potato) colors are also subject to manipulation using these techniques. Flower colors, of course, are comm nly very dependent on the activity f the flavonoid pathway genes, and thus are especially sensitive to the absolute and relative levels f expression of the flavonoid biosynthetic pathway genes. Ornamental plants and flowers are valuable commercially, and thus are typical targets of th methods herein described. Creation and selection of new coloration schemes are particularly valuable in the ornamental flower bearing plants such as chrysanthemums, carnations, roses, gerberas, lilies, geraniums, poinsettias and petunias.

Transformation

The transformation of plants in accordance with the invention may be carried out in essentially any of the various ways known to those skilled in the art of plant molecular biology. See, in general, *Methods in Enzymology* Vol. 153 ("Recombinant DNA Part D") 1987, Wu and Grossman Eds., Academic Press, incorporated herein by reference. As used herein, the term transformation means alteration of the genotype of a host plant by the introduction of a nucleic acid sequence. The nucleic acid sequence need not necessarily originate from a different source, but it will, at some point, have been external to the cell into which it is to be introduced.

In one embodiment, the foreign nucleic acid is mechanically transferred by microinjection directly into plant cells by use of micropipettes. Alternatively, the foreign nucleic acid may be transferred into the plant cell by using polyethylene glycol. This forms a precipitation complex with the genetic material that is taken up by the cell (Paszkowski et al., (1984) *EMBO J.* 3:2717-22).

In another embodiment of this invention, the introduced gene may be introduced into the plant cells by electroporation (Fromm et al., (1985) "Expression of Genes Transferred into Monocot and Dicot Plant Cells by Electroporation," *Proc. Natl Acad. Sci. U.S.A.* 82:5824, which is incorporated herein by reference). In this technique, plant protoplasts are electroporated in the presence of plasmids or nucleic acids containing the relevant genetic construct. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and form a plant callus. Selection of the transformed plant cells with the transformed gene can be accomplished using phenotypic markers.

Cauliflower mosaic virus (CaMV) may also be used as a vector for introducing the foreign nucleic acid into plant cells (Hohn et al., (1982) "Molecular Biology of Plant Tumors," Academic Press, New York, pp.549-560; Howell, U.S. Pat. No. 4,407,956). CaMV viral DNA genome is inserted into a parent bacterial plasmid creating a recombinant DNA molecule which can be propagated in bacteria. After cloning, the recombinant plasmid again may be cloned and further modified by introduction of the desired DNA sequence into the unique restriction site of the linker. The modified viral portion of the recombinant plasmid is then excised from the parent bacterial plasmid, and used to inoculate the plant cells or plants.

Another method of introduction of nucleic acid segments is high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads r particles, or on the surface (Klein et al., (1987) *Nature* 327:70-73). Although typically nly a

single introduction of a new nucleic acid segment is required, this method particularly provides f r multiple introductions.

A preferred method f introducing the nucleic acid segments into plant cells is to infect a plant cell, an explant, a meristem or a seed with *Agrobacterium tumefaciens* transformed with the segment. Under appropriate conditions known in the art, the transformed plant cells are grown to form shoots, roots, and develop further into plants. The nucleic acid segments can be introduced into appropriate plant cells, for example, by means of the Ti plasmid of *Agrobacterium tumefaciens*. The Ti plasmid is transmitted to plant cells upon infection by *Agrobacterium tumefaciens*, and is stably integrated into the plant genome (Horsch et al., (1984) "Inheritance of Functional Foreign Genes in Plants," *Science*, 233:496–498; Fraley et al., (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80:4803).

Ti plasmids contain two regions essential for the production of transformed cells. One of these, named transfer DNA (T DNA), induces tumor formation. The other, termed virulent region, is essential for the introduction of the T DNA into plants. The transfer DNA region, which transfers to the plant genome, can be increased in size by the insertion of the foreign nucleic acid sequence without its transferring ability being affected. By removing the tumor-causing genes so that they no longer interfere, the modified Ti plasmid can then be used as a vector for the transfer of the gene constructs of the invention into an appropriate plant cell, such being a "disabled Ti vector".

All plant cells which can be transformed by Agrobacterium and whole plants regenerated from the transformed cells can also be transformed according to the invention so as to produce transformed whole plants which contain the transferred foreign nucleic acid sequence.

There are presently at least three different ways to transform plant cells with Agrobacterium:

(1) co-cultivation of Agrobacterium with cultured isolated protoplasts,

(2) transformation of cells or tissues with Agrobacterium, or

(3) transformation of seeds, apices or meristems with Agrobacterium.

Method (1) requires an established culture system that allows culturing protoplasts and plant regeneration from cultured protoplasts.

Method (2) requires (a) that the plant cells or tissues can be transformed by Agrobacterium and (b) that the transformed cells or tissues can be induced to regenerate into whole plants.

Method (3) requires micropropagation.

In the binary system, to have infection, two plasmids are needed: a T-DNA containing plasmid and a vir plasmid. Any one of a number of T-DNA containing plasmids can be used, the only requirement is that one be able to select independently for each of the two plasmids.

After transformation of the plant cell or plant, those plant cells or plants transformed by the Ti plasmid so that the desired DNA segment is integrated can be selected by an appropriate phenotypic marker. These phen typic markers include, but are not limited to, antibiotic resistance, herbicide resistance r visual observation. Other phenotypic markers are known in the art and may be used in this invention.

All plants from which protoplasts can be isolated and cultured to give wh le regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred foreign gene. Some suitable plants include, for example, species from the genera Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna. Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Hyoscyamus, Lycopersicon, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Ciohorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Hererocallis, Nemesia, Pelargonium, Panicum, Pennisetum, Ranunculus. Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Lolium, Zea, Triticum, Sorghum, and Datura.

It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major cereal crop species, sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Limited knowledge presently exists on whether all of these plants can be transformed by Agrobacterium. Species which are a natural plant host for Agrobacterium may be transformable in vitro. Although monocotyledonous plants, and in particular, cereals and grasses, are not natural hosts to Agrobacterium, work to transform them using Agrobacterium has also been carried out (Hooykas-Van Slogteren et al., (1984) *Nature* 311:763–764). Additional plant genera that may be transformed by Agrobacterium include Chrysanthemum, Dianthus, Gerbera, Euphorbia, Pelaronium, Ipomoea, Passiflora, Cyclamen, Malus, Prunus, Rosa, Rubus, Populus, Santalum, Allium, Lilium, Narcissus, Ananas, Arachis, Phaseolus and Pisum.

Regeneration

Normally, regeneration will be involved in obtaining a whole plant from the transformation process. The term "transgenote" refers to the immediate product of the transformation process and to resultant whole transgenic plants.

The term "regeneration" as used herein, means growing a whole plant from a plant cell, a group of plant cells, a plant part or a plant piece (e.g. from a protoplast, callus, or tissue part).

Plant regeneration from cultural protoplasts is described in Evans et al., "Protoplasts Isolation and Culture," *Handbook of Plant Cell Cultures* 1:124–176 (MacMillan Publishing Co. New York 1983); M. R. Davey, "Recent Developments in the Culture and Regeneration of Plant Protoplasts," *Protoplasts*, (1983)—Lecture Proceedings, pp.12–29, (Birkhauser, Basal 1983); P. J. Dale, "Protoplast Culture and Plant Regeneration of Cereals and Other Recalcitrant Crops," *Protoplasts* (1983)—Lecture Proceedings, pp. 31–41, (Birkhauser, Basel 1983); and H. Binding, "Regeneration of Plants," *Plant Protoplasts*, pp.21–73, (CRC Press, Boca Raton 1985).

Regeneration from protoplasts varies from species to species of plants, but generally a suspension of transformed protoplasts containing copies of the exogenous sequence is first made. In certain species embryo formation can then be induced from the protoplast suspension, to the stage of ripening and germination as natural embryos. The culture media will generally contain various amino acids and h rmones, such as auxin and cytokinins. It is sometimes advantageous to add glutamic acid and proline to the medium, especially for such

species as corn and alfalfa. Shoots and roots n rmally develop simultaneously. Efficient regeneration will depend on the medium, on the genotype, and n the history f the culture. If these three variables are controlled, then regeneration is fully reproducible and repeatable.

Regeneration also occurs from plant callus, explants, organs or parts. Transformation can be performed in the context of organ or plant part regeneration. *Methods in Enzymology*, supra: also *Methods in Enzymology*, Vol. 118; and Klee et al., (1987) *Annual Review of Plant Physiology*, 38:467–486.

In vegetatively propagated crops, the mature transgenic plants are propagated by the taking of cuttings or by tissue culture techniques to produce multiple identical plants for trialling, such as testing for production characteristics. Selection of desirable transgenotes is made and new varieties are obtained thereby, and propagated vegetatively for commercial sale.

In seed propagated crops, the mature transgenic plants are self crossed to produce a homozygous inbred plant. The inbred plant produces seed containing the gene for the newly introduced foreign gene activity level. These seeds can be grown to produce plants that would produce the selected phenotype.

The inbreds according to this invention can be used to develop new hybrids. In this method a selected inbred line is crossed with another inbred line to produce the hybrid.

Parts obtained from the regenerated plant, such as flowers, seeds, leaves, branches, fruit, and the like are covered by the invention, provided that these parts comprise cells which have been so transformed. Progeny and variants, and mutants of the regenerated plants are also included within the scope of this invention, provided that these parts comprise the introduced DNA sequences. Progeny and variants, and mutants of the regenerated plants are also included within the scope of this invention.

Vectors

Selection of an appropriate vector is relatively simple, as the constraints are minimal. The apparent minimal traits of the vector are that the desired nucleic acid sequence be introduced in a relatively intact state. Thus, any vector which will produce a plant carrying the introduced DNA sequence should be sufficient. Also, any vector which will introduce a substantially intact RNA which can ultimately be converted into a stably maintained DNA sequence should be acceptable.

Even a naked piece of DNA would be expected to be able to confer the properties of this invention, though at low efficiency. The decision as to whether to use a vector, or which vector to use, will be guided by the method of transformation selected.

If naked nucleic acid introduction methods are chosen, then the vector need be no more than the minimal nucleic acid sequences necessary to confer the desired traits, without the need for additional other sequences. Thus, the possible vectors include the Ti plasmid vectors, shuttle vectors designed merely to maximally yield high numbers of copies, episomal vectors containing minimal sequences necessary for ultimate replication nce transformation has occurred, transposon vectors, hom logous recombination vectors, minichromosome vectors, and viral vectors, including the possibility f RNA forms f the gene sequences. The selection of vectors and methods to construct them are commonly known t persons f ordinary skill in the art and are described in general technical references (Methods in Enzymol gy, supra).

However, any additi nal attached vector sequences which will confer resistance to degradation of the nucleic acid fragment to be introduced, which assists in the process of genomic integration or provides a means to easily select for those cells or plants which are actually, in fact, transformed are advantageous and greatly decrease the difficulty of selecting useable transgenotes.

Selection

Selection of transgenotes for further study will typically be based upon a visual assay, such as observing color changes (e.g., a white flower, variable pigment production, and uniform color pattern on flowers or irregular patterns), but may involve biochemical assays of either enzyme activity or product quantitation. Transgenotes will be grown into plants bearing the plant part of interest and the gene activities will be monitored, such as by visual appearance (for flavonoid genes) or biochemical assays (Northern blots, see, Maniatis (below); Western blots, see, Ausubel (below); enzyme assays and flavonoid compound assays, including spectroscopy, see, Harborne et al., (Eds.), (1975) *The Flavonoids*, Vols. 1 and 2, [Acad. Press]). Appropriate plants will be selected and further evaluated.

The following experimental section is offered by way of example and not by limitation.

EXPERIMENTAL

In general, preparation of plasmid DNA, restriction enzyme digestion, agarose gel electrophoresis of DNA, Southern blots, DNA ligation and bacterial transformation were carried out using standard methods. (Maniatis et al., *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory (1982), referred to herein as "Maniatis" and hereby incorporated by reference.) Western blots and other standard molecular biology techniques are also described in Ausubel et al., (1987) *Current Protocols in Molecular Biology*, Vols. 1 and 2, and hereby incorporated by reference.

EXAMPLE 1

Plant Transformation Procedures

In Examples 1–4, reagent materials are commercially available, unless otherwise specified. Enzymes used in the cloning procedures are available from commercial sources. All restriction endonuclease reactions are carried out according to manufacturer instructions. Unless otherwise specified, the reaction conditions for other reactions are standard conditions used in the art, as described, for example, in Maniatis. Luria (L) agar and Minimal A (MinA) agar and broth are described in J. H. Miller, *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, New York (1972) (referred to herein as "Miller" and hereby incorporated by reference). Transformations of competent *Escherichia coli* strain DH-1 were performed according to Maniatis. Plasmid DNA was prepared by alkaline extraction according to Maniatis (referred to herein as "mini-prep DNA" or "mini-prep technique"). Site specific oligonucleotide mutagenesis was carried out as described in Geisselsoder, et al., (1987) *BioTechniques* 5:(8), 786–791, except *E. coli* strain BW313 (dut,ung) was used t produce uracil—containing single stranded DNA, in vitro synthesized double stranded DNA was transf rmed

into *E. coli* strain DH-1, and the Klen w fragment f DNA polymerase I was used for the second strand synthesis which was incubated vernight at room temperature.

Antibiotics are abbreviated as follows: Ap is ampicillin, Km is kanamycin, Rif is rifampicin and Tc is tetracycline. Micrograms are referred to herein as ug and milliliters are referred to as ml. Microliters are referred to as ul.

Creation of Agrobacterium Tumefaciens LBA4404/p5972 and p7506

The plasmid pcPE1 was obtained from H. Reif, Max Planck Institut, Koln. This plasmid contained a nearly full length *Petunia hybrida* chalcone synthase cDNA clone as an EcoRI fragment. (For the sequence of a chalcone synthetase gene, see, Niesbach-Klosgen et al. (1987) *J. Mol. Evolution* 26:213–225, which is incorporated herein by reference.) The construction strategy that led to the construction of the binary vectors p5972 and p7506 which were used for the reintroduction and expression of this chalcone synthase gene into target plants is shown in FIG. 1. The plasmids shown in this strategy are labeled with only the relevant restriction sites used either in the construction procedure or discussed in the text. The plasmid numbers in the middle of the circles in the figure are the actual number designations given to the plasmids in the construction strategy. The plasmid number in the lower bottom of the circles refers to the cloning vector that gave rise to the relevant clones. For example the first plasmid listed in the strategy is pcPE1. This clone resulted from a ligation of an EcoRI fragment into the commercially available cloning vector pUC18, so pcPE1 is listed in the middle of the circle and pUC18 is listed in the bottom. The restriction enzymes listed by the drawn circles indicate which enzymes were used to digest the plasmids and an arrow indicates that a ligation reaction took place. Antibiotic resistant genes that were used to select the clones are indicated inside the circles.

The EcoRI fragment containing the complete coding sequence for chalcone synthase protein was recloned into the EcoRI site of plasmid pUC119 (Viera and Messing, *In Methods in Enzymology*, 153(2) eds. Wu and Grossman, San Diego, pp.3–11, 1987) by digesting both plasmids with EcoRI, ligation, and transformation into competent *E. coli* strain DH-1. A plasmid which contained pUC119 and the EcoRI chalcone synthase fragment was identified by restriction mapping and was designated plasmid p6010. Plasmid p6010 was transformed into competent *E. coli* strain BW313 and single stranded DNA containing uracil was isolated (Viera and Messing, ibid.). A 26 base synthetic primer composed of the sequence, 5'-CTTTTTTCTAGTTAACCATGGTGACT-3', and a 24 base synthetic primer composed of the sequence, 5'-CTACTTAGTGGATCCGGCTTATAT-3', were synthesized on an Applied Biosystems 381A DNA synthesizer using the solid phase phosphotriester method. The 26 base primer was used to introduce two new restriction sites, HpaI and NcoI, at the beginning of the coding sequence. The NcoI site overlapped the ATG start codon of the chalcone synthase and would be used f r promoter fusions later in the construction strategy. The 24 base primer was used to introduce a BamHI site that overlapped the TAG translation stop codon and would be used later in the construction strategy to fuse the chalcone synthase gene to a poly-adenylation signal sequence. In vitro synthesized d uble stranded DNA (dsDNA) was synthesized using the tw primers and transformed into competent *E. coli* strain DH-1. Ampicillin resistant colonies were screened using mini-prep DNA f r new HpaI and NcoI restriction sites that mapped at the beginning of the gene and a BamHI site that mapped at the end of the gene. The plasmid that fulfilled this and further mapping criteria was designated as plasmid p5571.

The next step in the construction strategy served to fuse a 35S Cauliflower Mosaic Virus (herein called CaMV 35S promoter in the text and p35S in the figure) to the beginning of the chalcone synthase coding sequence and a poly-adenylation signal sequence to the end of the coding sequence. Plasmid pJJ2104 is described in Harpster et al., (1988) *Mol. Gen. Genet.* 212:182–190 which is incorporated herein by reference, and was used as the source of the CaMV 35S promoter and poly-adenylation signal sequence. This plasmid has a modified CaMV 35S promoter contained within a BglII and NcoI fragment. The CaMV 35S promoter within plasmid pJJ2104 is fused to the untranslated leader sequence of the photosynthetic 22L chlorophyll a/b binding protein (here in called "Cab22L") to increase transcriptional efficiency (see above reference). The poly-adenylation signal sequence is from the nopaline synthase gene (Depicker et al., (1982) *Mol. Appl. Genet.* 1(6):561–573), and is contained within a BamHI and HindIII fragment in the plasmid pJJ2104.

Plasmid p5571 DNA was isolated and cleaved to completion with BamHI and then cleaved with NcoI under conditions to give a partial, incomplete digestion of the DNA because a second NcoI site lies within the chalcone synthase coding sequence. The DNA was subjected to electrophoresis through 0.5% low melt agarose in standard Tris-Acetate EDTA buffer (described in Maniatis) with ethidium bromide at 0.5 ug/ml incorporated into the agarose. The gel was examined briefly under medium length ultraviolet light using a transilluminator (wave length 312 nm) and a band corresponding to the length (approximately 1200 base pairs) of the chalcone synthase coding sequence was excised from the gel. The gel fragment was weighed to determine the volume and brought to 0.3M sodium acetate. The equilibrated agarose was heated to 65 degrees for the 10 minutes and then extracted with an equal volume of phenol saturated with 0.1M Tris-HCl, pH 8. The aqueous phase was removed and extracted twice with a chloroform-isoamyl alcohol (24:1) mixture and the DNA was then precipitated from the aqueous solution. (All the above techniques were used according to standard conditions as described in Maniatis.) This eluted fragment was combined with plasmid pJJ2104 which was cleaved to completion with NcoI and BamHI and a ligation reaction was set up and incubated for one hour at room temperature and the reaction products subsequently transformed into competent *E. coli* strain DH-1. Plasmid DNA was screened with restriction enzymes to identify the appropriate plasmid containing the 35S CaMV promoter, the chalcone synthase coding sequence, and the poly-adenylation signal sequence. The plasmid p5634 was identified and subjected to restriction digestion to confirm that the plasmid was the correct one.

Two different binary vectors were used in the construction strategy. Plasmids pJJ3942 and pAGS502 are both based on the broad host range cloning vector pRK290 (Ditta et al., (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77:7347–7451) and contain a Neomycin Phosphotrans-

ferase II coding sequence fused at the 5' end to a nopaline synthase promoter and at the 3' end t an octopine synthase poly-adenylation signal sequence between the left and right TDNA borders (van den Elzen et al., (1985) *Plant Mol. Biol.* 5:141–154). Plasmid pAGS502 contains a polylinker with cl ning sites for BamHI, XbaI, HindIII, XhoI, EcoRI and HpaI for insertion of fragments near the TDNA right border. Plasmid pJJ3942 contains HindIII, BamHI, and HpaI as unique cloning sites near the right border. An enhancer-like sequence from the 35S CaMV promoter is contained between the BamHI and HpaI sites. This fragment spans the sequences between positions −45 and −200 of the 35S CaMV promoter to give approximately 200 bases of sequence upstream from the TATAA box.

The entire 35S CaMV promoter, the chalcone synthase coding sequence and the nos poly-adenylation signal sequence are contained within a BglII and HindIII fragment in plasmid p5634. Plasmid p5634 was digested to completion with BglII and HindIII. Two different binary vectors, pJJ3942 and pAGS502, were digested to completion with BamHI and HindIII and each was used in separate ligation reactions with plasmid p5634 digested with BglII and HindIII. The 5'-overhangs generated by the enzymes BglII and BamHI can ligated together but not recleaved by either enzyme. The ligation reactions were transformed into competent *E. coli* strain DH-1 and tetracycline resistant colonies were isolated. DNA was isolated using the mini-prep technique and screened with the appropriate restriction enzymes to isolate pJJ3942 and pAGS502 derivative plasmids that accepted the BglII and HindIII fragment. Further restriction digestions were performed to confirm the identity of the resulting plasmids. The ligation product of pJJ3942 containing the insert was designated as plasmid p5972 and the ligation product of pAGS502 containing the insert was named plasmid p7506.

Plasmids p5972 and p7506 were mobilized (transferred) separately to *Agrobacterium tumefaciens* strain LBA4404 (Ooms et al., (1981) *Gene* 14:33–50). A triparental mating procedure using *E. coli* strain DH-1 carrying p5972 or p7506 (both tetracycline resistant), *E. coli* HB101 carrying plasmid pRK2013 (kanamycin resistant) (Ditta et al., (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77:7347–7351) and *A. tumefaciens* strain LBA4404 (rifampicin resistant) was set up. The two *E. coli* strains were grown up overnight on L agar (see Miller) containing the appropriate antibiotics. The *A. tumefaciens* was grown up overnight in MinA broth (see Miller) with no selection. One ml of the *A. tumefaciens* culture was pipetted into a sterile microcentrifuge tube and spun in a microcentrifuge for 2 minutes to pellet the cells. The supernatant was removed and 100 ul of fresh MinA broth was added to resuspend the pellet. A small amount of the *E. coli* cells from each of the overnight cultures was scraped off the petri dish and spread together onto a fresh L agar plate (no antibiotics). The amount of area covered by the cells was approximately 2 cm square. Each amount of *E. coli* cells was approximately equal to the amount of *A. tumefaciens* cells that was collected from 1 ml of culture. The 100 ul of resuspended *A. tumefaciens* cells was added on top of the spread *E. coli* cells and mixed to form a conjugation patch. This petri dish was incubated overnight at room temperature.

On the following day approximately one-fourth of the cells was rem ved from the conjugation patch and these cells were streaked for single colonies using an L agar plate containing 100 ug/ml rifampicin and i.2 ug/ml tetracycline. The procedure was repeated four times and resulted in all of the conjugation patch streaked ont four separate plates. These plates were incubated in the dark at room temperature until colonies begin to appear (approximately 3–5 days). Isolated colonies were streaked for single colonies on MinA agar plates containing 1.2 ug/ml tetracycline. The plates were incubated for two days at 28° C. A petri dish containing MinA agar supplemented with 1.2 ug/ml tetracycline was divided into eight equal parts of a circle and eight well isolated single colonies were streaked individually onto sections of the petri dish. This plate was grown up overnight at 28° C. Three-fourths of the cells from each of the eight sections were removed from the agar using a sterile toothpick and the DNA isolated from these cells using the mini-prep technique. Each of the DNAs from these eight preparations was transformed individually into competent *E. coli* strain DH-1 and tetracycline resistant colonies were isolated. One colony from each *E coli* transformation was grown up and the DNA isolated using the mini-prep technique. The DNA was subjected to restriction enzyme analyses to confirm that the DNA was the original binary clone that was transferred to *A. tumefaciens* LBA4404 via the triparental mating.

Plant Transformation

Petunia hybrida varieties: Pink Cascade was obtained from Dr. Michael Reid, Dept. of Environmental Horticulture, University of California, Davis; R18 and V26 were obtained from Dr. Anton Gerats, Dept. of Genetics, Free University, Amsterdam. *Petunia hybrida* plants were grown from surface-sterilized seed on sterile solidified agar medium of 1/10 the concentration of medium MS of Murashige supplemented with 0.5% sucrose. After germination, seedling tops were excised by cutting in the hypocotyl region and transferred to MS with 3% sucrose. Plants were maintained at 28° C. under "cool white" fluorescent light at 4–5000 lux, 16h/day.

About six weeks after planting (day 0), leaves were excised, cut with a scalpel blade into pieces about 5 mm square and inoculated with *A. tumefaciens* that had been grown overnight in MinA medium supplemented to 0.2% glucose (medium described by J. H. Miller (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, New York) and adjusted to 0.1–0.2 A_{550} units. Inoculated leaf pieces were placed on incubation medium [basal MS medium (MS+3% sucrose +B5 vitamins)+75–100uM acetosyringone, 1 mg benzyladenine (BA) per liter, and 0.2 mg indoleacetic acid (IAA)/l] for two days in a sterile transfer hood at room temperature (approx 22° C.). On day 2, 25–30 ml of liquid basal MS medium+cefotaxime (500 mg/l) was added to the plates. Plates were then swirled at 70–100 rpm for 30–60 min. Leaf pieces were transferred with the upper epidermis facing up on selection medium (basal MS+BA (1 mg/l), IAA 0.2 mg/l, vancomycin (100 mg/l)). The plates were sealed with parafilm and incubated at 24° C. under moderate light (3000–5000 lux). On day 14, leaf pieces were transferred to fresh selection medium. On day 28, calli were excised from leaf pieces and transferred to fresh selection medium and shoots were excised and transferred t h rm neless medium [basal MS+vancomycin (100 mg/l) and kanamycin (100 mg/l)]. On day 42 and following, shoots were excised from calli and transferred t hormoneless

medium. After shoot elongation, shoots were excised and dipped in naphthalene acetic acid (NAA) (0.1 mg/l) for root development. After rooting, plantlets were transplanted to soil and grown in a greenhouse.

The chimeric CHS gene in p5972 and p7506 was introduced into several varieties of Petunia: (1) a hybrid variety called "Pink Cascade", (2) an inbred, R18, and (3) an inbred, V26. (See. examples 2–4)

EXAMPLE 2

Novel Derivatives of Pink Cascade Petunia

Pink Cascade produces solid pink flowers. Leaf explants from the Pink Cascade variety were transformed with p5972. Six whole plants (CS18201 through 18206) were produced. All had novel flowers. CS18201, 18203, and 18206 gave pure white petal limbs, petal tubes and anthers. CS18202 and 18205 gave flowers with a color pattern: pink wedges at the outer margin of and in the center of petal limbs with the rest of the flower pure white (some flowers on this plant were solid pure white, other flowers had this pattern). CS18204 flowers were a light, blotchy pink.

Progeny of the cross V26 x CS18202 included: 12 plants with the color of V26 x Pink Cascade and 6 plants with novel color patterns similar to the patterns of CS18202, but with smaller pigmented sectors on lower petals than upper petals in some progeny. Thus, the production of novel color patterns by the introduced gene is heritable, but the pattern itself may vary among progeny (because Pink Cascade is a hybrid variety, the progeny are genetically heterogenous).

EXAMPLE 3

Novel Derivatives of the R18 Variety Transformed with p5972

Variety R18 produces solid, pale pink flowers. Cells from the R18 variety were transformed with p5972 to produce 14 plants. Nine plants produced flowers with the normal light pink color of R18 flowers. The flowers of five plants had novel patterns. One plant gave pink radial stripes on a solid white background. Another gave mainly pure white flowers, but one flower had some pink wedge similar to wedges on CS18202. The third gave occasional white wedges at petal junctions on a solid pink background. The fourth gave a mixture of pure white flowers and pink with white radial striations (a star-like pattern). The last gave white flowers with pink wedges at the outer margins of petals.

EXAMPLE 4

Novel Derivatives of the V26 Variety Transformed with p5972 and p7506

Variety V26 produces solid, deep violet flowers. Cells from the V26 variety were transformed with p5972 to produce 37 plants. Twenty-eight plants produced flowers colored the same as the V26 parent. Seven plants had flowers with novel patterns; two plants produced pure white flowers. Three plants had pigmented wedges at petal margins similar to CS18202. One had mostly pure white flowers, but some flowers had single, small (3 mm) white spots. One plant gave flowers having a beautiful "Cossack dancer" pattern, i.e., a modified radial, star-like pattern. Two plants gave flowers with a somewhat irregular, blotchy pattern of white and purple patches; these flowers, though irregular, looked somewhat like the dancer pattern.

Cells from the V26 variety were transformed with p7506 to produce 20 plants. Seventeen plants produced flowers colored the same as the V26 parent. Three plants produced flowers with color patterns. One plant produced flowers with occasional small white spots. One plant produced one flower with a white tube, while the other flowers were similar to the V26 parent. One plant had flowers which exhibited nearly randomly distributed but sharply defined blotches.

Several transgenotes with white or patterned flowers were crossed to V26. The progeny of a white transgenote produced violet and white flowers in approximately a 1:1 ratio, as expected for a single gene. The progeny of plants with patterned flowers were also patterned or sometimes pure white. Not all plants were identical in flower color intensity or pattern. The penetrance of the flower color phenotypes was complete in the progeny populations of some transgenotes and incomplete in others, i.e., the segregation ratio of solid violet to patterned or solid white was significantly greater than 1 (for incomplete penetrance).

ANALYSES ON PETUNIA HYBRIDA TRANSGENIC PLANTS

RNA Analyses: The steady state levels of messenger RNA from both wild type plants (i.e., with an endogenous chalcone synthase gene) and transgenic plants (i.e., with an introduced 35S CaMV driven chalcone synthase gene) were analyzed in petunia petal using RNase protection analyses (protocol is titled *RNA Transcription,* available from Stratagene, 11099 North Torrey Pines Road, La Jolla, Calif.).

Petals of six different developmental stages were first harvested from V26 petunia plants and from one transgenic V26 which had white flowers (plant #21838). These stages were defined according to total petal length, degree of pigmentation, and morphology:

Stage	Length	Pigmentation & Morphology
1	15 mm	no pigmentation (veins only)
2	30	slight flush of light purple around veins
3	40	definite pigmentation from outer surface
4	53	deeper pigmentation, fully extended, still closed
5	58	fully pigmented, just starting to open
6	nd*	freshly mature, fully expanded

*nd; not determined

RNA was isolated from the above described developmental stages for both wild type V26 and transgenic plant #21838. One flower of each stage except stage 1 was sufficient tissue for extracting RNA. For stage 1, eight to ten flowers were combined for the procedure. Petal tissue was frozen in liquid nitrogen and ground to a fine powder in liquid nitrogen using a pre-chilled mortar and pestle. The tissue was added to 1 ml phenol saturated with 0.1M Tris-HCl, pH 7.5 and 4 ml buffer (100 mM NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1% SDS) and the contents mixed. One ml of chloroform: isoamyl alcohol mixture (24:1) was added and the contents mixed again. The aqueous phase was transferred to a clean tube and extracted a second time with fresh chloroform: isoamyl alcohol mixture. The aqueous phase was transferred to a clean tube, an equal volume of 4M lithium acetate was added and the contents of the tube were placed on ice for three hours. The RNA was pelleted by centrifugation and the supernatant was removed. The pellet was dissolved in sterile

water, the solution was brought to 0.3M sodium acetate and 2.5 v lumes f ethanol were added to precipitate the RNA. The RNA pellet was dissolved in 100 ul f sterile water and the concentration of RNA was determined spectroph tometrically.

Five ug of RNA was used for each protection assay. A 160 nucleotide, radi labelled anti-sense cab22L-CHS RNA was transcribed in vitro, used as a probe in the protection assays and annealed to the petal RNAs, all as described in the Stratagene protocol. After incubation with single strand specific ribonucleases RNase A and RNase T1, two different protected fragments will remain, a 94 nucleotide fragment representing the endogenous CHS mRNA and a 160 nucleotide fragment representing the introduced chalcone synthase transcript.

An autoradiogram of the RNase protection assays for all six stages of wild-type V26 petunia petal RNAs showed that the chalcone synthase protection fragment was most abundant in stage 3 and stage 4. From this experiment it was determined that the endogenous CHS mRNA is present in petals at all developmental stages examined, gradually increasing in abundance up to stage 4 and then declining to almost undetectable levels in the mature petal.

RNase protection assays on transgenic plant #21838 showed that protection fragments for both the endogenous and the introduced chalcone synthase were present. The relative levels of the endogenous chalcone synthase message followed a similar developmental profile as seen for wild-type plants; however, the overall message levels were substantially reduced, as observed by visual inspection, in each stage from the levels seen in the wild-type V26 plant. In contrast to the wild type chalcone synthase message, the chalcone synthase message from the introduced 35S CaMV promoter was present at a fairly constant low level throughout each developmental time point. This result demonstrates that the introduced CHS gene had the effect of vastly depressing the steady state level of endogenous CHS mRNA.

Protein Analyses

Antibodies were raised in a rabbit against chalcone synthase by injecting the rabbit with a fusion protein made in *E. coli.* This fusion protein consisted of wild type beta-galactosidase gene with the entire coding sequence of chalcone synthase ligated in-frame to the 3' end of beta-galactosidase (Ruther and Müller-Hill (1983) *EMBO J.* 2:(10):1791–1794). Immune antiserum from the rabbit was used in Western analyses to evaluate wild-type and transformed petunia petals. Western analysis were carried out according to manufacturer's instructions using the Proto Blot system from Promega Biotec; but similar techniques are described in Ausubel et al., (supra).

Protein extracts were prepared from purple and white segregants (described above). The same developmental stages as described above were used. Petal tissue was frozen in liquid nitrogen and then ground to a fine powder using a mortar and pestle. The frozen powder was transferred to a glass tissue homogenizer and extraction buffer (50 mM sodium phosphate buffer pH 7.0, 1 mM dithiothreotol, 0.1% Triton X-100, 1 mM EDTA) was used to homogenize the tissue. Protease inhibitors were added to final concentrations of 1 mM phenylmethysulfonyl fluoride and 0.2 mM Leupeptin. Cell debris was removed by centrifugation and the protein content f the supernatant was determined using the Bradford assay (Bradford (1976) *Anal. Biochem.* 72:248–254). Seventeen ug f pr tein for each sample was loaded nto an 8% polyacrylamide SDS gel.

Western analysis of purple flowered progeny sh wed that CHS protein was present in the petal extract f all devel pmental stages. The amount of CHS protein appeared to be the same for stages 1 and 2, increased in stage 3, stayed approximately the same for stages 3, 4 and 5 and decreased slightly for stage 6. Western analysis of white flowered progeny showed that, in comparison to the purple flowered progeny, barely detectable CHS protein was seen in stage 1 and appeared to be more reduced in stages 2, 3, 4, 5 and 6. These analyses showed that while CHS protein could be easily detected in protein extracts from purple flowered progeny, in protein extracts of white flowered progeny CHS protein was reduced to levels where it was barely detectable.

TLC Analysis of Flavonoids

Thin layer chromatography (TLC) was done to compare flavonoid synthesis in white versus purple flowers from progeny of the cross #21838×V26. Mitchell petunia flowers were used as a negative control for anthocyanin synthesis and a positive control for flavonoid synthesis. Flowers of three different lengths were used as follows: Mitchell (33 mm, 43 mm, 65 mm), purple (33 mm, 42 mm, 55 mm) and white (35 mm, 49 mm, 57 mm). The tubes were assayed separately from the limbs. Tissue was added to 1.0 ml of 2N HCl, allowed to stand for 2 hours at room temperature and then hydrolyzed for 20 minutes at 100° C. The supernatant was transferred to a clean tube and 200 ul of isoamyl alcohol was added. The samples were vortexed for at least five seconds and the two phases allowed to separate. Samples were spotted onto a cellulose TLC plate in four separate applications with drying between applications. Two identical plates were set up and run in two different solvent systems; acetic acid/36% HCl/water (30:3:10) and isopropanol/2N HCl (1:1). The two systems discriminate between the anthocyanins. The purple progeny flowers produced both anthocyanins and flavonols. The Mitchell flowers produced flavonols and little or no anthocyanins. The white progeny flowers produced little or no anthocyanins and little or no flavonols of the type produced by the normal purple flowers.

TLC Analysis of Anthocyanin Precursors

The marked reduction in anthocyanin accumulation in the white progeny flowers described above indicates that there is a block in the pathway for anthocyanin biosynthesis. Analysis was performed to determine at which step of the anthocyanin biosynthetic pathway the biochemical block occurred. Extracts were prepared as described above from flowers of 22838 and 21841 plants, as well as the V26 parent. As is described in Table I, the level of caffeic acid was increased significantly in 21838 and 21841 relative to the level present in V26. The level of coumaric acid was relatively unchanged.

Coumaric acid is expected to be present in V26 flowers both in the precursor pool and bound to anthocyanin (based on the Rt genotype of V26 (Heller and Forkmann, "Biosynthesis", in *The Flavonoids,* Harborne, J. B., ed., Chapman and Hall Ltd., London, 1988, pp. 399–425, and references therein). The presence f anthocyanin-bound coumaric acid complicates the analysis in terms f coumaric acid precursor levels. To deter-

mine whether the coumaric acid detected in TLC analysis above derives from the precursor r anthocyanin-bound form, a number of control petunia lines were similarly analyzed f r precursors. In particular, levels f coumaric acid in the petunia line R27, which lacks the Rt gene necessary for coumaric acid linkage to anthocyanins, were compared to levels of coumaric acid in V26. As can be seen in FIG. 1, there was no detectable coumaric acid in extracts from line R27 (lanes 15 and 16), despite the presence of anthocyanin at levels comparable to V26 (band with Rf of 0.86).

To confirm that coumaric acid is generally not detected in petunia lines lacking the anthocyanin-bound form, extracts were prepared from flowers of lines W80 (mutated in the dihydroflavonol-4-reductase gene), W85 (mutated in the dihydroflavonol-4-reductase gene), W37 (mutated in the flavanone-3-hydroxylase gene) and R77 (mutated in both flavonoid 3'-hydroxylase and 3', 5'-hydroxylase genes) (Heller and Forkmann, and references therein). These mutants W37, W80, W85 and R77 produce little or no anthocyanin, but the mutational block is not at chalcone synthase (Heller and forkmann, and references therein). No coumaric acid was detected in corollas from any of these lines. This indicates that the coumaric acid detected in line V26 is not from the precursor pool, but rather is bound to anthocyanin.

The two white lines, 21838 and 21841, showed the accumulation of significantly more caffeic acid and coumaric acid as a precursor (i.e., not bound to anthocyanin) than did the parent V26, consistent with a biochemical block at the chalcone synthase step.

TABLE I

TLC Analysis of Petunia Lines

LINE	FLOWER	PLANT	ANTHOCYANIN	CAFFEIC	COUMARIC
V26	Purple	control	+	+	+
V26 (CHS)	White	21838	−	++	+
V26 (CHS)	White	21841	−	++	+
R27	Red	control	+	+	−
R77	Red	control	+/−	+	−
W37	White	control	−	+	−
W80	White	control	−	+	−
W85	White	control	−	+	−

The lines are described in the test. Symbols: −, none detected; +, detected at a level equivalent to that of line V26; ++, detected at a level significantly above that of line V26.

EXAMPLE 5

Genomic Chalcone Synthase (CHS)

Experimental procedures in Examples 5 and 6 were performed at Vrije Universiteit according to the teachings of the present invention (see, van der Krol, Mur, Beld, Mol and Stuitje, "Flavonoid Genes in *Petunia hybrida:* Addition of a limited number of gene copies may lead to a collapse in gene expression," Plant Cell, Vol. 2 pages 291-299 (1990)). Unless stated otherwise, standard methods were carried out in accordance with protocols as found in Maniatis and Bernard Perbal (1988), *A Practical Guide to Molecular Cloning,* John Wiley and Sons, New York.

A. Preparation of Transformants

VIP76 and VIP106 were derived from the *Petunia hybrida* V30 CHS gene A genomic clone VIP17, as described in Koes et al., (1986) *Nucl. Acids Res.* 14:5229-5239. Specifically, VIP76 was constructed by cloning a 7.2 kb Ec RI-SalI fragment fr m VIP17, containing the CHS gene A with a 0.8 kb 5' promoter region, into the EcoRI-SalI site of the binary vector BIN19 (Bevan et al., (1984) *Nuc. Acids Res.* 12:8711-8721). VIP106 was constructed by cl ning a 8.0 kb Xba I fragment from VIP17, containing the CHS gene A with a 4.5 kb 5' promoter region, into the Xba I site of BIN19.

For transformation of *Petunia hybrida* VR plants (van der Krol et al., (1988) *Nature* 333:866-869), the binary vectors with the different gene constructs were mobilized into *Agrobacterium tumefaciens* strain LBA 4404 using standard triparental mating techniques (Ditta et al., (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77:7347-7351). Transformants were obtained through the standard leaf-disc transformation method (Horsch et al., (1985) *Science* 227:1229-1231). Transformed *Petunia hybrida* VR plants were grown in a greenhouse.

B. Description of Transformants

VIP76 contains 0.8 kb of the 5' promoter region of the CHS A gene, 2.5 kb of the CHS A gene exon and intron sequences and 3.9 kb of the CHS A gene 3' flanking DNA. VIP106 contains 4.6 kb of the 5' promoter region of the CHS A gene, 2.5 kb of CHS A gene exon and intron sequences, and 0.5 o kb of the CHS A gene 3' flanking DNA. Both genomic CHS clones were introduced into petunia VR plants. Of the twenty plants regenerated which contain VIP106, one plant showed an evenly reduced pigmentation of the corolla. One of the 15 plants containing the VIP76 construct showed white flowers with small pigmented sections.

C. Messenger RNA Data

RNA was extracted from flowerbuds (stage 1-6 as defined by Koes et al., (1989) *Plant Mol. Biol.* 12:213-226). To average out small differences in developmental stage of the flowerbuds, five flowerbuds taken from one plant were pooled per stage. Nucleic acid isolation was performed as described by Koes et al., (1987) *Plant Mol. Bio.* 10:375-385.

Transcripts from VIP106 and VIP76 contain CHS gene A sequences that derive from the petunia strain V30. They can be distinguished from endogenous petunia VR CHS gene A transcripts by primer extension experiments. Primer extension experiments using ^{32}P-labelled oligomer EL-4 (5'-dGATCAACACAGTTT-GTAGG-3') and 5 ug floral RNA (stage 4) were performed by annealing the oligomer and RNA in 10 ul 100 mM Tris HCl, 20 mM $MgCl_2$ and 100 mM KCl for 16 hours at 30° C. Then 10 ul of 1 mM dGTP, 1 mM dATP, 1 mM dTTP, 1 mM dCTP containing 5U of reverse transcriptase were added, and the solution was incubated for 30 minutes at 37° C. followed by 30 minutes at 43° C. Extension products were recovered by phenol chloroform (1:1) extraction and ethanol precipitation, and then visualized on 6% sequencing gels using sequence ladders of CHS gene A primed with the same oligonucleotide as a marker. Extension of primer EL-4 hybridizing to the first exon of CHS gene A results in one major fragment (176 nucleotides) for V30 CHS gene A mRNA, and two major fragments and one minor fragment for the VR CHS gene A mRNAs (186 and 189 nucleotides and 181 nucleotides respectively). Normal pigmented transformants containing VIP106 and VIP76 have normal expression levels of VR CHS mRNAs in pigmented floral tissue. VIP78 transformant

76-1A, in which flower pigmentation was inhibited, sh wed a lower than n rmal level of VR CHS mRNAs.

EXAMPLE 6

Dihydroflav nol Reductase (DFR)

A. Preparation f Transf rmants

VIP178 was constructed using a 1.3 kb EcoRI DFR cDNA fragment from clone lambda DFR-A1, which was isolated as follows. A lambda gt11 cDNA library, as described in van Tunen et al., (1988), *EMBO J.* 7:1257-1263, was prepared according to Young and David, (1983) *Proc. Natl. Acad. Sci. U.S.A.*, 80:1194-1198. The library, prepared from corolla tissue of flowerbuds of *Petunia hybrida* line R27 (Van Tunen et al., (1988) *EMBO J.* 7:1257-1263) was screened with a 1.4 kb EcoRI/BamHI fragment of the pallida gene (Martin et al., (1985) *EMBO J.* 4:1625-1630). Fourteen clones hybridizing to this probe were purified. Thirteen of these contained an insert of nearly the same size (1.5 kb), whereas one clone contained an insert of around 1.6 kb. Sequence analysis of the 5' and 3' ends of both types of clones revealed that the smaller cDNA insert (lambda DFR-A1) was nearly full-length, whereas the larger cDNA insert had an extended 3' region and was shorter at the 5' end. Sequence analysis showed no differences between the clones. All fourteen clones showed an identical hybridization behavior, indicating that these clones represent transcripts of the same gene. The EcoRI sticky-ends were filled-in using the Klenow fragment of DNA polymerase I and dNTPs. This blunt-ended fragment was cloned into the HincII site of M13 mp7 so it could subsequently be isolated as BamHI fragment. This was cloned into the BamHI site of VIP103 (van der Krol et al., (1988) *Nature* 333:866-869) from which the CHS sequences had been removed. The clone which contains the BamHI DFR fragment in sense orientation to the CaMV 35-S promoter was called VIP178.

For transformation of *Petunia hybrida* VR plants, the binary vectors with the different gene constructs were mobilized into *Agrobacterium tumefaciens* strain LBA 4404 using standard triparental mating techniques (Ditta et al. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77:7347-7351). Transformants were obtained through the standard leaf-disc transformation method (Horsch et al. (1985) *Science* 227:1229-1231). Transformed petunia hybrida VR plants were grown in the greenhouse.

B. Description of Transformants

VIP178 contains the CaMV 35-S promoter, a full-length cDNA copy of the petunia R27 DFR gene and a nopaline synthase 3' tail fragment. The DFR gene construct (VIP178) was introduced into petunia VR. Of the 25 plants regenerated, 19 plants show a flower phenotype indistinguishable from that of untransformed petunia VR plants. However, six plants show a reduction in flower pigmentation in sector or ring patterns. The degree of pigmentation varies among different flowers of the same plant. For instance, flowers on transformant 178-16 are either completely pigmented or have a white corolla.

C. Messenger RNA Data

RNA was extracted from flowerbuds (stage 1-6 as defined by Koes et al., (1989) *Plant Mol. Biol.* 12:213-226). T average ut small differences in devel- pmental stage f the fl werbuds, five flowerbuds taken fr m one plant were pooled per stage. Nucleic acid isolation was perf rmed as described by Koes et al., (1987) *Plant Mol. Biol.* 10:375-385. The ^{32}P-labeled DFR antisense RNA was synthesized fr m the vector pTZ18U in which the full-length DFR cDNA EcoRI fragment was cloned. This pr be is completely protected by the transcripts f the VIP178, while RNase protection with petunia VR DFR RNA results in multiple fragments due to sequence divergence between the DFR genes of petunia lines R27 and VR.

The RNase protection experiments permitted distinction between the DFR mRNA from the sense DFR gene construct (VIP178, transcript completely protected by the R27-DFR probe) and the petunia VR DFR mRNAs (two subfragments upon RNase protection with the R27-DFR probe). For transformants 178-16 and 178-17, the DFR mRNA steady-state level in flowerbud stage 4 resulting from expression of the transgene, together with that of the endogenous DFR, genes was analyzed. In 178-17 a high expression of the DFR transgene was observed relative to that of endogenous DFR gene(s). In pigmented floral tissue of 178-16 the same effect was noted. In contrast, a severe reduction in both the endogenous as well as the transgene DFR mRNA steady-state levels was observed in white flora tissue of transformant 178-16.

EXAMPLE 7

Novel Derivatives of the Chrysanthemum Variety Moneymaker Transformed with a Chrysanthemum CHS Sense Construct

A. Isolation of a Chrysanthemum CHS cDNA

A cDNA library was constructed in the phage λgt10 made from the commercially available chrysanthemum variety "Debonair" according to the standard methods of Hynh et al. (DNA Cloning Techniques: A Practical Approach, David Glover, ed., IRL Press, Oxford 1984, pp. 49-78, which is incorporated herein by reference). The library consisted of 5 petri plates, totalling over 200,000 recombinant lambda phage plaques. The mRNA for cDNA synthesis was prepared from flower petal tissues, as described in example 4. A hybridization probe was synthesized according to standard techniques based on highly conserved CHS sequences as described in Niesbach-Klosgen et al. (J. Molecular Evolution, 26, 213-25, 1987, which is incorporated herein by reference). The sequence of the oligonucleotide is:

5'CCTCCAGCAAAGCAACCCTGTTG-
GTACATCATG 3'

Bacteriophage lambda plaques were selected on the basis of hybridization to the oligonucleotide above using standard plaque lift methodology. Hybridizing phage clones were grown up for the isolation of DNA. DNA was prepared and then digested with restriction enzyme EcoRI for subcloning into the phagemid pBLUESCRIPT KS (Strategene Cloning systems, La Jolla, Calif.). These phagemid clones were retested for hybridization with the oligonucleotide. Clones which had sequences giving a hybridization signal were subjected to sequence analysis. Using primers based on pBLUESCRIPT KS sequence, cDNA clone sequence was obtained up to about 150 bp of each end. For ne such clone, 68% sequence homology was f und at the 5' end of the cDNA clone, and 71% sequence hom l- gy was found at the 3' end f the cDNA clone. The

amino acid homol gy ranged fr m 78% t 91% from the amino terminus to the carboxy terminus. This sequence homology is consistent with that found by Niesback-Klosgen et al. The cDNA clone with these sequences was theref re deemed to correspond to the transcript fr m an authentic chrysanthemum CHS gene.

B. Construction of Vectors for Introduction Into Chrysanthemum

The CHS cDNA sequence was modified to introduce an NcoI site at the start codon and a BamHI site just downstream of the stop codon, essentially as described in Example 1. The resulting plasmid was then treated with restriction enzymes NcoI and BamHI, and the fragment was ligated with an NcoI-BamHI fragment containing a promoter sequence and a polyadenylation signal sequence (plasmid pJJ2104, see Example 1). The reaction products were transformed subsequently into *E. coli* strain DH5α (Woodcock et al., 1989, Nuc Acids Res., 17, 3469–78), and selected for resistance to ampicillin. Bacterial colonies were picked, plasmid DNA isolated, and the resulting plasmid screened with restriction enzymes to identify the appropriate plasmid containing the 35S CaMV promoter, the chrysanthemum chalcon synthase coding sequence, and the poly-adenylation signal sequence. A plasmid was isolated with the appropriate characteristics.

To facilitate cloning in the next step, an SphI site was replaced with an EcoRI site in the plasmid above. An SphI site at the junction between the polyadenylation signal sequence and the cloning vector was replaced with an EcoRI site by first digesting with SphI, removing the 4 base-pair overhang with T4 DNA polymerase, and then ligating to EcoRI linkers. After digesting with EcoRI, a ligation reaction was performed prior to transformation into DH5α. Bacterial colonies were picked and tested for loss of the SphI site and the presence of an EcoRI site at the same location. The resulting plasmid was designated pFLG9336.

Plasmid pAGS802 is a binary vector used for the introduction of the chrysanthemum CHS construct contained in pFLG9336. The binary vector contains left and right border regions from *Agrobacterium tumefaciens* T-DNA (van den Elzen et al., Plant Mol. Biol., 5, 149–54, 1985), an NPTII coding sequence (Berg at al., Proc. Natl. Acad. Sci. U.S.A., 72, 3628–32, 1975) fused at the 5' end to the CaMV 35S promoter and at the 3' end to an octopine synthase poly-adenylation signal (see Example 1), and the lacZα multiple cloning region of pUCI19. The plasmid harboring this T-DNA construct also has a minimal C01E1 replicon, obtained as the AlwnI-PvuII 180 bp fragment of pBR322 (Bolivar et al., Gene, 2, 95–113, 1977), and a *Pseudomonas aeruginosa* replicon, pVSI, obtained as a SacII-BamHI fragment (Stanisich et al., J. Bacteriol., 129, 1227–33, 1977). Plasmid pAGS802 has a unique EcoRI site in the lac multiple cloning region, into which was inserted the EcoRI fragment carrying the chrysanthemum CHS fusion construct of pFLG9336. After ligation reaction products were transformed into *E. coli* strain MJ83 (Yanish-Perron et al., 1985, Gene, 33, 103–99) and selected for gentamycin resistance, bacterial colonies were grown up for plasmid DNA isolation. Plasmid DNA was screened for the presence of the EcoRI fragment containing the CHS coding sequence construct in pAGS802, giving a plasmid designated pFLG4716 (see FIG. 3).

C. Introduction f the Chrysanthemum CHS Gene Into Chrysanthemum Variety Moneymaker

Plasmid pFLG4716 was introduced int *E. coli* strain JM83, and then m bilized into *Agrobacterium tumefaciens* strain LBA4404 as described in Example 1, with gentamycin at 10 ug/ml being used as selecti n agent rather than tetracycline at 1.2 ug/ml. Individual transconjugant colonies were purified, and a single colony selected for chrysanthemum transformation. The colony was grown as described in Example 1, DNA was isolated, and plasmid DNA transformed into *E. coli* strain SK1592 (Yanish-Perron et al., 1985, Gene, 33, 103–99). The DNA was subjected to restriction enzyme analysis to confirm that the DNA retained the original structure of the binary plasmid that was transferred to *A. tumefaciens* LBA4404 via the triparental mating.

Alternatively, transformation was performed with *A. tumefaciens* strain LBA4404 containing either plasmid pAGS802 or PAGS190. Plasmid pAGS802 was described above; pASG190 is a similar derivative, except that it does not contain the lacZα fragment, and it contains a gene construct encoding hygromycin resistance (Gritz and Davies, 1983, Gene, 25, 179–88). The gene construct has the hygromycin phosphotransferase gene (HPT) coding sequence and the same transcriptional start and termination signals used for the NPTII gene.

Chrysanthemum (*Dendranthema grandiflora*) variety Moneymaker was obtained from Fides BV, De Lier, The Netherlands. It was maintained and propagated by placing nodal segments (7–10 mm) into Magenta GA-7 cubes (Magenta Co., Chicago) with 80 ml of an MS-based medium (MS major salts, minor salts and iron, Murashige et al., 1962, *Phys. Planta.*, 15, 473–97), B-5 vitamins (Gamborg et al., 1968, *Exp. Cell Res.*, 50, 148–51), 3% sucrose, 3 mM 2- (N-Morpholino) ethanesulfonic acid (MES), pH 5.7 (with KOH) which has been solidified with 0.7% Tissue Culture agar (Hazelton Biologics, In.). The in vitro plants were grown in a controlled environment room at 24.C under 16 hours illumination/day at 150–200 foot-candles of plant spectrum fluorescent illumination (Agro Lite, Philips).

Partially to fully expanded leaves and petioles from the in vitro plants were removed and placed in a few mls of sterile distilled water in a 150×20 mm petri plate. Two 4 mm wide sections, each of petiole and leaf, were made. The water was removed and the explants were then inoculated with *Agrobacvterium tumefaciens* strain LBA4404 containing plasmid pFLG4716 (as described above) which had been grown on Minimal A medium (Davis et al., 1950, J. Bacteriol., 60, 17–28). The explants were then removed from the excess inoculum and placed on 7.0 cm sterile Whatman #1 filter paper circles on top of incubation medium plates, specifically MS-based medium as above with 0.2% Gelrite (Scott Laboratories, Inc.), 2.0 mg/l 6-BA (6-benzylaminopurine), 1.0 mg/l NAA (2-naphthaleneacetic acid), 200 μM 3', 5'-dimethoxy- 4'-hydroxy-acetophenone (acetosyringone). The incubation medium plates were incubated at room temperature (around 22° C.) for 3 days.

Following the incubation period, explants were transferred to selection medium plates, specifically, MS-based medium solidified with 0.2% Gelrite, containing 2.9 mg/l 6-BA, 1.0 mg/l NAA, 500 mg/l Geopen (carbenicillin, Pfizer, Roerig) and 100 mg/l kanamycin sulfate. Plates were placed in clear plastic boxes and put in a controlled environment room at 24° C. with 150–200 foot-candles illumination for 16 h urs/day. After 6–7

weeks, regenerated shoots were removed from the explants and placed on rooting medium, specifically, MS-based medium solidified with 0.2% Gelrite with 500 mg/l Geopen and 50 mg/l kanamycin sulfate to confirm transformation. These shoots were evaluated after 16 days for rooting in the presence of kanamycin sulfate. Rooted plantlets were transplanted to soil and grown in a greenhouse to flowering.

nia as described in a prior example above). This data is summarized in Table II.

Accumulation of caffeic and/or coumaric acid is indicative of a block at the chalcone synthase step in the pathway for flavonoid biosynthesis. This indicates that plants A, B and C, derived from transformation with a CHS gene construct, had either a complete or partial biochemical block at the chalcone synthase step.

TABLE II

	TLC Analysis of Transgenic Chrysanthemum Plants				
GENE CONSTRUCT	FLOWER COLOR	PLANT	ANTHOCYNIN	CAFFEIC	COUMARIC
—	pale pink	Control	+	+	+
pFLG4716 (CHS)	pale pink	—	+	+	+
pFLG4716 (CHS)	extremely light pink	C	−	+	++
pFLG4716 (CHS)	white	A,B	−	++	++
pAGS802 (−)	pale pink	Control	+	+	+
pAGS190 (−)	pale pink	Control	+	+	+
pAGS190 (−)	extremely light pink	D	−	+	+

Plants A, B, C and D are described above. Symbols: −, none detected; + detected at a level equivalent to that of Moneymaker (parent); ++, detected at a level significantly above that of parent (at least 30% more).

D. Characteristics of Chrysanthemum CH Transgenic Chrysanthemums

Chrysanthemum variety Moneymaker has a pale pink color. Explants were transformed with pFLG4716 as described above to produce 133 CHS-transgenic plants. An additional 134 plants were transformed with a binary vector lacking the CHS construct. Plants were brought to flowering by subjecting them to short day conditions, created by covering the greenhouse area with black cloth for 12–13 hours each day for 80 days. Of the 267 plants 8 plants had flowers which were slightly lighter pink than the flowers of the parent, 9 plants had flowers which were yellow, and 16 had flowers which were darker pink than the flowers of the parent.

Of the 133 plants resulting from pFLG4716 transformation, the ray flowers of 2 plants were pure white (A, B), and the ray flowers of 1 plant were extremely light pink (C). No plants with white color were obtained from Moneymaker transformed with vector alone. One extremely light colored plant was obtained in a control transformation (D).

Several plants with unaltered flower color, both CHS-transformed white flower plants (A, B), the CHS-transformed extremely light colored flowering plant (C), and the control-transformed extremely light flowering control plant (D) were analyzed by thin layer chromatography for flavonoids. Ray flowers were separated from disc flowers for this analysis. The flowers from one composite flower were placed into 2.0 ml of 2N HCl, and then heated for 20 minutes at 100° C. An aliquot of the hydrolysate was transferred to an eppendorf tube and extracted with 0.2 volumes of isoamyl alcohol. Samples were run on cellulose thin layer plates as described in Example 4.

None of the plants produced flavonols; the plants with unaltered colors produced the same pattern of anthocyanins as the original Moneymaker. Plants A, B, C and D produced no detectable anthocyanins, consistent with the lack of color in the flowers. Plants A, B and C produced enhanced levels of bands with mobilities corresponding to that of caffeic acid and coumaric acid. In contrast, plant D did not accumulate additional caffeic or coumaric acids. There was no evidence of coumaric acid bound to anthocyanins (unlike V26 petu-

EXAMPLE 8

Acetolactate Synthase (ALS)

In this example, *Nicotiana tabacum* carrying two loci, SuRA and SuRB (Chaleff et al. (1986) Molecular Strategies for Crop Protection, UCLA Symposium on Molecular and Cellular Biology (Arntzen and Ryan, eds.), 48:15–425), which each encode the enzyme acetolactate synthase (ALS), was used in a transformation study. ALS is involved in the first common step in the biosynthesis of branched chain amino acids in plants. ALS is the site of, and is sensitive to, the action of sulfonylurea (SU) herbicides; La Rossa et al., (1984) *J. Biol. Chem.* 259:8753–8757. A single amino acid replacement in the ALS protein results in a herbicide resistant form of the enzyme (Lee et al. (1988) *EMBO Journal* 7:1241–1248); this effect can be mediated by mutation of either the SuRA or SuRB locus.

In this study, as detailed below, tobacco cells were transformed with DNA containing a deleted form of the ALS gene which would, if full length, encode a resistant form of the ALS enzyme. Homologous recombination led to several clones which were resistant (as calli) to SU as the result of replacement of the sensitive form of the ALS gene at the SuRB locus by deleted and resistant form of the ALS gene. These clones had varying numbers (one to three) of copies of the introduced fragment in addition to the substituted SuRB locus. Upon regeneration into a plant, one of the clones (containing three extra copies of the fragment) was found to have regained its sensitivity to SU. That is, the sensitivity phenotype of the SuRA locus (not homologously recombined) was manifested, and the resistance phenotype of the SuRB locus which was the site of the homologous recombination (and for which there were additional copies in the genome of the mutant fragment) was masked or suppressed. This is evidence for the repression of the ALS resistance phenotype in accordance with the method of the invention.

Experimental details and results are set forth below.

Method logy

Construction f the Targeting Plasmid

The plasmid pAGS157 contains a tobacco mutant ALS gene with three nucleotide replacements (586, 587,

1719) which give rise to the two amino acid replacements (196:pro→ala and 573:trp→leu); either one f these amino acid replacements confers herbicide resistance to the encoded ALS enzyme (Lee et al. (1988) *EMBO Journal* 7:1241–1248); pAGS167 carries the corresponding wild-type gene. These plasmids were used to construct the targeting plasmid. In vitro mutagenesis (Kunkel (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82:488–492) was used to introduce a diagnostic ApaI site (AGAGCACAC$_{1753}$→AGGGCCCAG$_{1753}$') 25 bp downstream from the 573:trp→leu mutant site in pAGS157, to form the plasmid pAGS175. This new site did not change the amino acid sequence of the protein. Subsequently, a 1.0 kb NcoI-HpaI fragment from pAGS175, containing the 196:pro→ala mutation, was replaced by the equivalent fragment from pAGS167, containing the wild-type sequence. The resulting plasmid, pAGS177, contains the single 573:trp→leu mutation and the introduced ApaI site. pAGS177 was ligated between the T-DNA borders at the BamHI site of the binary vector pAGS140 (Dean et al. (1988) *Nucleic Acids Res.* 16:7601–7617) to give the plasmid pAGS180BV. This plasmid was used as a positive control. Plant cells arising from transformation with this plasmid are resistant to hygromycin, kanamycin, and chlorsulfuron. An inactive form of the herbicide resistant ALS gene with a 1.5 kb deletion removing the promoter and the 5' 510 bp of the ALS coding region was constructed as follows. The plasmid pAGS177 was digested with ClaI and PstI and the single-strand overhangs removed by treatment with T4 DNA polymerase and dNTPs. The 1.86 kb fragment containing the 573:trp→leu mutation and the novel ApaI site was purified from an agarose gel and cloned into the HpaI site between the T-DNA borders in the binary vector pJJ2525. The resulting plasmid was named pAGS182BV and was used for the experiments. It has a total of 1.86 kb of homology to the SuRB ALS gene. The binary vectors pAGS180BV and pAGS182BV were conjugated into the Agrobacterium strain LBA4404 (Hoekema et al. (1983) *Nature* 303:179–180) by triparental mating with *E. coli* strain HB101/pRK2013 (Figurski and Helinski (1979) *Proc. Natl. Acad. U.S.A.* 76:1648–1652).

Agrobacterium Co-cultivation with Tobacco Protoplasts

The protoplasts were obtained from leaves of *N. tabaccum* var Wisconsin 38 plants propagated by nodal culture. Co-cultivation of protoplast-derived tobacco colonies with *A. tumefaciens* was carried out as described previously (van der Elzen et al. (1985) *Plant Mol. Biol.* 5:149–154). Selection for kanamycin or chlorsulfuron resistance was initiated 8 days after protoplast isolation. Co-cultured plant cells were plated on a medium containing either kanamycin (50 μg/ml) or chlorsulfuron (2 ng/ml), and this medium was replaced twice weekly. Plants were regenerated by placing colonies on the medium of Murashige and Skoog (Murashige and Skoog (1962) *Physiol. Plant* 15:473–497) containing 0.1 μg/ml NAA, 1.0 μg/ml BAP and 10 ng/ml chlorsulfuron.

Southern Analysis of Genomic DNA

Genomic DNA was isolated from selected herbicide resistant callus and shoot tissue grown in Magenta boxes, or from leaves of regenerated plants grown in the greenhouse (Dooner et al. (1985) *Mol. Gen. Genet.* 200:240–246). Approximately 10 ug f DNA was digested with 50 units of the appropriate restriction endonuclease for 2–4 hours at 37° C. Hybridization and washing were carried ut as described (Lee et al. (1988) *EMBO Journal* 7:1241–1248) t high stringency. Probes were prepared either by using ribopr be vect rs (Promega, Madison, Wis.) or by random priming (Boehringer Mannheim, Indianapolis, Ind.) as recommended by the manufacturers.

Segregation Analysis

Transformed plants were self-pollinated, and the resulting seed were plated on MS media (Murashige and Skoog (1962) *Physiol. Plant* 15:473–497) containing either 200 ug/ml kanamycin or 50 ng/ml chlorsulfuron. Seedlings were scored after 10 to 14 days as either sensitive or resistant depending on their ability to form roots.

RESULTS

Targeting Strategy

The targeting DNA consists of a mutant ALS gene from the SuRB locus conferring herbicide resistance which has o been inactivated by deletion of the 5' coding sequences and promoter. The resistance mutation is a trp-leu change at amino acid 573; an ApaI restriction site was introduced 25 bp downstream of this mutation to mark the targeting DNA (this sequence alteration did not change the amino acid sequence of the encoded ALS protein).

Since the coding sequences between the SuRA and SuRB genes are highly conserved (Lee et al. (1988) *EMBO Journal* 7:1241–1248), it is possible that recombination could occur at each locus, however, the products of recombination at either locus can be differentiated by their restriction digest pattern.

Transformation with the Nonfunctional Mutant ALS Gene Yields Chlorsulfuron Resistant Colonies

Tobacco (cv. Wisconsin 38) protoplasts were co-cultivated with *A. tumefaciens* carrying pAGS180BV (the control plasmid) or pAGS182BV (which carries the deleted fragment from the mutant ALS gene). Following co-cultivation, protoplasts were divided and plated on chlorsulfuron (2 ng/ml) or kanamycin (50 ug/ml). Transformation with pAGS182BV produced seven chlorsulfuron resistant clones. Transformed calli and the regenerated plants derived from these calli were maintained on 10 ng/ml chlorsulfuron until transfer to the greenhouse. Each herbicide resistant transformant was tested for kanamycin resistance by transferring callus pieces to plates containing 200 ug kanamycin/ml and monitoring survival. Of the 7 chlorsulfuron resistant transformants selected, 2 (HR11 and HR15) were also kanamycin resistant.

Three Chlorsulfuron Resistant Clones Result from Homologous Recombination

To determine if the chlorsulfuron resistant clones resulted from homologous recombination or spontaneous mutation, Southern hybridizations were conducted on genomic DNA from the 7 chlorsulfuron resistant transformants. As a probe, an EcoRI-HindIII fragment was used from the SuRA gene which spans the introduced mutations and extends 450 bp in the 5' and 750 bp in the 3' direction.

After genomic DNA is digested with NcoI and ApaI, the two endogenous ALS genes, SuRA and SuRB, give rise to hybridizing bands of 4.7 kb and 2.0 kb. Homolo-

gous recombination between the pAGS182BV DNA and the ALS genes at SuRA or SuRB should result in the appearance of two new bands of 1.2 kb and 0.8 kb.

In each instance the hybridization pattern for DNA isolated from three independent chlorsulfuron resistant transformants (HR11, HR14 and HR15) digested with NcoI and ApaI showed the bands predicted if homologous recombination had occurred. The other four chlorsulfuron resistant transformants had hybridization patterns which were identical to untransformed tobacco (data not shown); hence, it is presumed they have arisen from spontaneous mutation at the SuRA or SuRB loci. The Southern hybridization patterns for DNA from HR11, HR14 and HR15 showed the 1.2 kb and 0.8 kb bands predicted to result from homologous recombination at the SuR loci. In DNA from HR11 and HR14, there appeared to be no ALS hybridizing bands other than those attributed to homologous recombination, suggesting additional T-DNA random insertion events had not occurred. In the plant HR15 three extra hybridizing ALS bands occurred, suggesting three additional insertions of T-DNA in this transformant.

Southern hybridization was also carried out following digestion of DNA with SpeI, which cuts in the genomic sequences flanking the SuR loci. When tobacco genomic DNA digested with SpeI and ApaI, the endogenous ALS genes should give rise to hybridizing bands of 10 kb and 3.4 kb. Homologous recombination between pAGS182BV DNA and SuRA should result in new 5.4 kb and 0.8 kb bands, while recombination at SuRB should result in hybridizing fragments of 2.6 kb and 0.8 kb.

Hybridization bands were analyzed for the genomic DNA isolated from regenerated plants after digestion with SpeI and ApaI. The 2.6 kb band predicted following homologous recombination at SuRB appeared only in DNA from the putative recombinants HR11, HR14 and HR15. The other hybridizing fragments in these DNAs apparently represent the SpeI-ApaI fragments generated from the random insertion of the T-DNA carrying the targeting ALS gene into the genome. As expected, based on the NcoI-ApaI hybridization patterns, HR15 had at least two other hybridizing bands in the SpeI-ApaI digestions. An additional band was also seen in HR11 DNA, suggesting at least one T-DNA random insertion. These data are consistent with the observation that HR11 and HR15 are also kanamycin resistant, indicating that they contain a functional NPTII gene. The results from SpeI-ApaI hybridizations show that the novel ApaI restriction site marker in the targeting DNA is now linked to a restriction site outside the SuRB gene.

Chlorsulfuron Resistance Segregates as a Single Mendelian Gene in the Recombinant Plants HR11 and HR14

The results of genetic analysis of the transformants HR11, HR14 and HR15 are summarized below. The segregation of chlorsulfuron resistance to sensitivity in the progeny from self-pollinated HR14 was consistent with a ratio of 3:1 resistant:sensitive, indicating a single dominant locus for chlorsulfuron resistance. All seed derived from HR14 were kanamycin sensitive. Progeny from self-pollinated HR11 gave a ratio of chlorsulfuron resistance to sensitivity consistent with either a 3:1 r 2:1 ratio, indicating a single Mendelian locus with the possibility f partial penetrance r homozygous lethality. The segregati n of kanamycin resistance in these progeny was consistent with a single locus for kanamycin resistance. When grown on media containing both kanamycin and chlorsulfur n, these progeny segregated 9:7 resistant:sensitive. This suggested that the markers f r kanamycin resistance and chlorsulfuron resistance were not linked. Progeny from HR15 were n t chlorsulfuron resistant (0/500 seed). On kanamycin the ratio of resistant to sensitive progeny was consistent with a 63:1 ratio, indicating the presence of three independent loci for kanamycin resistance.

Overall, the observed phenotype of the HR15 plant is consistent with the introduced resistant fragment of the ALS gene, which replaced the sensitive form of the gene at the SuRB locus, being suppressed by the additional three other copies which occur (in a non-expressed state) in the genome. This suppression was not immediate upon the introduction of the resistant ALS genes since the callus from which HR15 was derived was herbicide resistant; however, with the ongoing cell division and plant regeneration, the herbicide resistant phenotype was suppressed.

EXAMPLE 9

Novel Derivatives of Plants with Modified Production of Plant Oils

The coding region of the stearyl desaturase (SD) gene is isolated in the same manner as described above in Example 1 for the chalcone synthase gene, and using information known in the art about lipid biosynthetic pathways at the genetic level (see, e.g., citations on same above). The SD gene is modified to contain isolated flanking restriction sites for introduction into the pJJ2104 plasmid. This provides a genetic construct having a modified 35S Cauliflower Mosaic Virus promoter in the correct orientation to produce a sense transcript having a poly-adenylation signal from the nopaline synthase gene. A modified photosynthetic 22L chlorophyll a/b leader sequence is also included as described above.

The new construct is checked to verify the correct orientation and vector content. A triparental mating, as described above in Example 1, is then performed using corresponding vector constructs and strains for the SD gene.

Plant transformation with *A. tumefaciens* clones is performed as described in Example 1. After transgenote cells are generated, whole plants are regenerated as described. Screening of plants exhibiting desired modified oil contents is performed by assaying the oil or lipid production of the transgenote plants on HPLC columns or other methods in accordance with standard procedures.

EXAMPLE 10

Novel Derivatives of Plants with Altered Production of Sugars or Carbohydrates

An isolated petunia starch synthase gene is substituted for the chalcone synthase gene in the procedure described in Examples 1 and 8. The procedure is performed essentially as described except for the screening. A standard iodine based starch assay or other assay for starch content of a biological sample is used to determine, for selection, the plant transgenotes exhibiting a modified starch content.

All publications are herein incorporated by reference to the same extent as if each individual publicati n r patent application was specifically and individually

indicated to be incorporated by reference. The invention now being fully described, it will be apparent to one of rdinary skill in the art that many changes and modifications can be made thereto without departing from the spirit r scope of the claims.

What is claimed is:

1. A method for producing a plant exhibiting at least one modified phenotypic trait by suppressing expression f an endogenous gene in the plant, said method comprising the steps of:

transforming plant cells with a polynucleotide to produce transgenote cells, the polynucleotide comprising a promoter operably linked to a DNA segment such that transcripts of the segment are produced in the sense orientation in the transgenote cells which segment transcripts have at least 65% sequence identity to transcripts of said endogenous gene and are effective to suppress expression of said endogenous gene in said plant cell;

growing plants from one or more of said transgenote cells, wherein production of mRNA encoded by the endogenous gene is reduced in one or more of the plants; and

selecting a plant exhibiting said modified phenotypic trait.

2. A method of claim 1, wherein said plant is a flowering plant.

3. A method of claim 1, wherein the segment encodes a full-length protein.

4. A method of claim 1, wherein the segment encodes a flavonoid metabolic pathway protein.

5. A method of claim 1, wherein the segment encodes a protein in the fatty acid biosynthetic pathway.

6. A method of claim 1, wherein the segment encodes a protein in the sugar to starch biosynthetic pathway.

7. A method of claim 1, wherein said promoter is a constitutive promoter.

8. A method of claim 1, wherein a transcript of the segment in the transgenote cells is identical to a transcript of the endogenous gene sequence in the transgenote cells.

9. A method of claim 1, wherein said polynucleotide is in a DNA vector.

10. A method for reducing production of a protein product of an endogenous gene in a plant, said protein affecting a phenotypic trait in the plant, said method comprising the steps of:

growing plants from a plant cell transformed with a DNA segment under operational control of a promoter, wherein transcripts of the DNA segment are in the sense orientation produced which are substantially homologous to transcripts of an endogenous gene encoding the protein, and wherein production of the protein is reduced; and

identifying a plant that exhibits a phenotypic change by screening the plants for modification of the phenotypic trait.

11. A method of claim 10, wherein said plant cells are transformed with a Ti plasmid vector comprising said DNA segment.

12. A method f claim 11, wherein said vector comprises a constitutive promoter.

13. A method f claim 10, wherein said promoter is heterologous to said plant.

14. A method of claim 10, wherein said segment encodes a flavonoid metabolic pathway gene.

15. A method of claim 10, wherein said plant is a dicotyledonous plant.

16. A method of claim 10, wherein said plant produces flowers.

17. A method of claim 10, wherein said plant is selected from the group consisting of the genus Petunia or chrysanthemum.

18. A method for modifying the coloration of chrysanthemum flowers comprising altering flower pigment production by transforming a chrysanthemum cell with a chalcone synthase gene or gene segment contained in a disabled Ti plasmid of *Agrobacterium tumefaciens*, wherein said chalcone synthase gene is operably linked to a promoter.

19. A plant of the genus chrysanthemum produced by the method of either of claims 1 or 10 exhibiting altered flower coloration, said plant comprising plant cells transformed with a sequence which is transcribed into an RNA sequence substantially homologous to an endogenous flavonoid biosynthetic pathway gene transcript.

20. A plant of claim 19, wherein said plant cells are transformed with a Ti plasmid vector.

21. A plant of claim 19, wherein said RNA sequence is at least about 500 bp in length and encodes a chalcone synthase protein fragment.

22. A method of claim 1, wherein said transcripts of the segment produced in the transgenote have greater than about 80% sequence identity to said transcripts of said endogenous gene.

23. A method of claim 1, wherein a plant exhibiting the modified phenotypic trait is selected visually.

24. A method of claim 1, wherein a plant is selected by identifying reduced levels of mRNA encoded by the endogenous gene.

25. A method according to claim 10, wherein the transcripts of the DNA segment correspond to a full length transcript of the endogenous gene.

26. A method of modifying a plant cell phenotype by reducing the level of mRNA encoded by an endogenous gene in the plant cell, said method comprising:

growing plant cells transformed with a polynucleotide comprising a transcribable DNA segment located downstream from an operably linked promoter, whereby the segment is transcribed in the cell in the sense orientation, wherein the transcribed DNA segment has a sequence complementary to the mRNA endoded by the endogenous gene and whereby transcription of the DNA segment is effective to reduce the level of mRNA transcribed by said endogenous gene; and

screening the cells for the modified phenotype.

27. A method according to claim 26, wherein the promoter is constitutive.

28. A method according to claim 26, wherein the endogenous gene is dihydroflavonal reductase.

29. A method according to claim 26, wherein the plant cell is a dicot plant cell.

30. A method according to claim 26, wherein the plant cell is a tobacco protoplast.

31. A method according to claim 26, wherein the transcribable segment encodes an herbicide resistant tobacco acetolactate synthase gene.

* * * * *

United States Patent [19]

Jorgensen et al.

[11] Patent Number: 5,034,323

[45] Date of Patent: Jul. 23, 1991

[54] **GENETIC ENGINEERING OF NOVEL PLANT PHENOTYPES**

[75] Inventors: **Richard A. Jorgensen; Carolyn A. Napoli**, both of Oakland, Calif.

[73] Assignee: **DNA Plant Technology Corporation**, Oakland, Calif.

[21] Appl. No.: **331,338**

[22] Filed: **Mar. 30, 1989**

[51] **Int. Cl.**[5] **C12N 15/29**; A01H 4/00

[52] **U.S. Cl.** **435/172.3**; 800/205; 800/DIG. 67; 536/27; 935/30; 935/35; 935/67

[58] **Field of Search** 800/1; 435/172.3, 317.1; 935/35, 64, 67, 30, 55; 536/27

[56] **References Cited**

PUBLICATIONS

Hahlbrock, "Flavonoids" in *The Biochemistry of Plants* (E. Conn, ed.) vol. 7 (1981) Academic Press, pp 425–456.

Reif, et al. (1985) Mol. Gen. Genet. 199: 208–215.

Meyer, et al. (1987) Nature 330: 677–678.

Horsch, et al. (1985) Science 227: 1229–1231.

Hall, et al. (1986) New Phytologist 103: 33–43.

Van der krol, et al. (Jun. 1988) Nature 333: 866–869.

Promega Catalog 1988/1989 Wisconsin U.S.A. p. 2.

Mol, et al. (1983) Mol. Gen. Genet. 192: 424–429.

Reeves et al., Genetics and Cytogenetics, Biol. Abst. 83(2): AB–454–455, 13846 (1986).

Krishnan et al., Plant Physiology, Biochemistry and Biophysics, Biol. Abstr. 82(7):AB–855, 67053 (1986).

Larsen, Carlsberg Res. Commun. 52:381–392 (1987).

Scherer and Knauf, Plant Molecular Biology 9:127–134 (1987).

G. Forkmann and B. Kuhn, (1979) *Planta* 144:189–192.

P. Dunsmuir et al., (1987) "Promoters and the Overexpression of Foreign Genes in Plants" in *Plant Biotechnology* (T. Mabry, Ed.) IC[2] Institute 107–119.

C. Martin et al., (1985) *EMBO J.* 4:1625–1630.

A. Van Tunen et al., (1988) *EMBO J.* 7:1257–1263.

R. Spribille and G. Forkmann, (1982) *Planta* 155:176–182.

R. Koes et al., (1987) *Plant Molecular Biology* 10:159–169.

R. Koes et al., (1986) *Nucelic Acids Research* 14:5229–5239.

U. Reimbold et al., (1983) *EMBO J.* 2:1801–1805.

F. Kreuzaler et al., (1983) *Proc. Natl. Acad. Sci. USA* 80:2591–2593.

U. Niesbach-Klösgene t al., (1987) *J. Mol. Evol.* 26:213–225.

H. Wiering and P. de Vlaming, "Genetics of Flower and Pollen Colors" in *Petunia, Monographs on Theoretical and Applied Genetics*, vol. 9 (K. C. Sink, Ed.), Springer Verlag, Berlin (1984).

C. Martin et al., (1987) "The Control of Floral Pigmentation in Antirrhinum Majus" in *Developmental Mutants in Higher Plants* (Thomas & Grierson, Eds.) Cambridge University Press.

J. Harborne, "Functions of Flavonoids in Plants" in Chemistry and Biochemistry of Plant Pigments (T. Goodwin, Ed.) 2d ed., vol. 1 (1976) Academic Press, London.

Primary Examiner—Elizabeth C. Weimar

Assistant Examiner—Che S. Chereskin

Attorney, Agent, or Firm—Townsend and Townsend

[57] **ABSTRACT**

Methods are provided for producing plants exhibiting one or more desired phenotypic traits. In particular, transgenotes are selected that comprise a DNA segment operably linked to a promoter, wherein transcription products of the segment are substantially homologous to corresponding transcripts of endogenous flavonoid biosynthetic pathway genes.

5 Claims, 2 Drawing Sheets

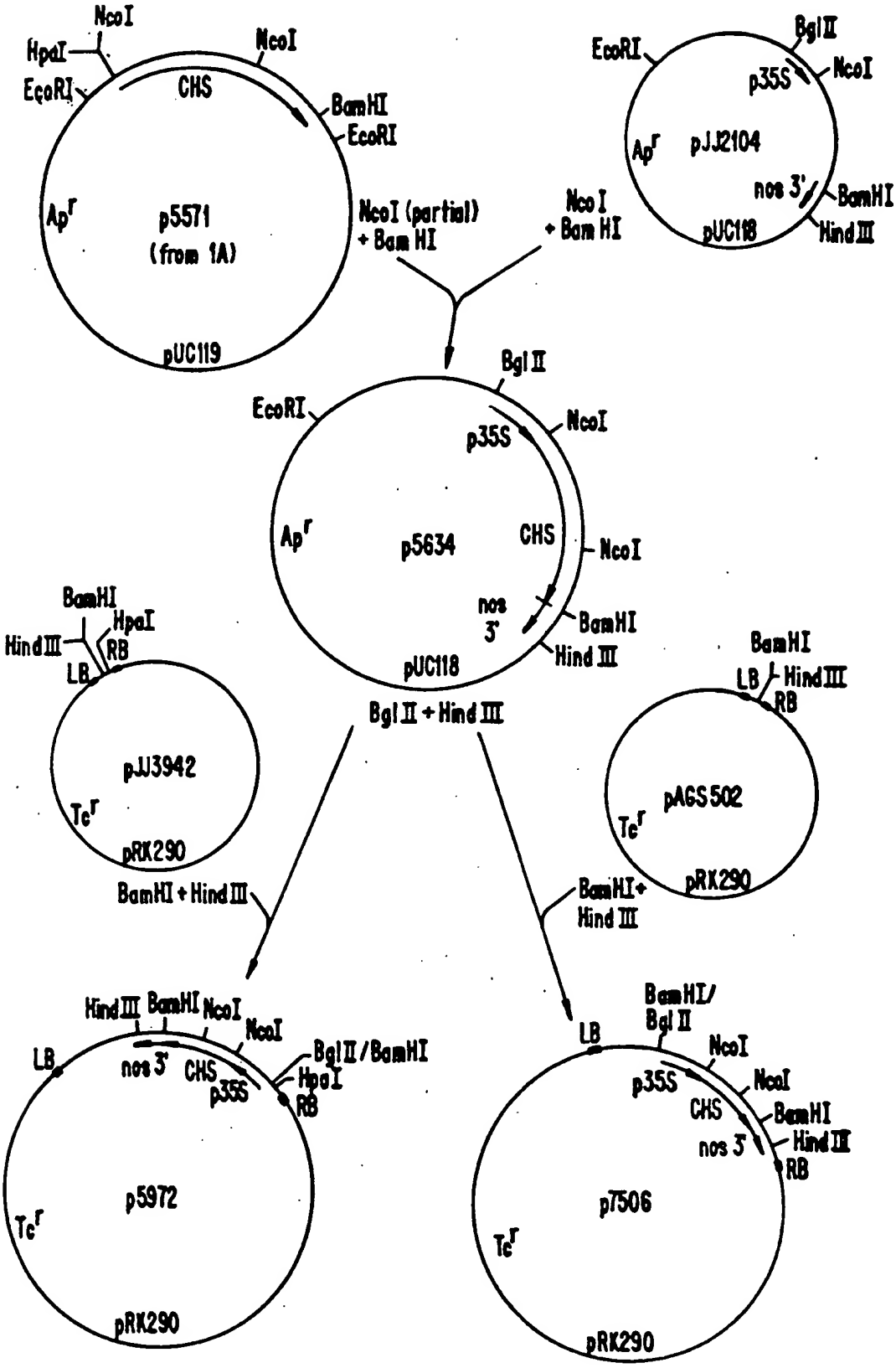

NcoI
HpaI
EcoRI
NcoI
CHS
BamHI
EcoRI
Ap^r
p5571
(from 1A)
pUC119
NcoI (partial)
+ Bam HI
NcoI
+ Bam HI
EcoRI
BglII
p35S
NcoI
Ap^r
pJJ2104
nos 3'
BamHI
pUC118
HindIII
BglII
EcoRI
NcoI
p35S
Ap^r
p5634
CHS
NcoI
nos
3'
BamHI
pUC118
HindIII
BglII + HindIII
BamHI
HpaI
HindIII
RB
LB
pJJ3942
Te^r
pRK290
BamHI + HindIII
BamHI
LB
HindIII
RB
pAGS502
Te^r
pRK290
BamHI+
HindIII
HindIII
BamHI
NcoI
NcoI
LB
nos 3'
CHS
p35S
BglII/BamHI
HpaI
RB
p5972
Te^r
pRK290
BamHI/
BglII
LB
NcoI
p35S
NcoI
CHS
BamHI
nos 3'
HindIII
RB
p7506
Te^r
pRK290

FIG._1B.

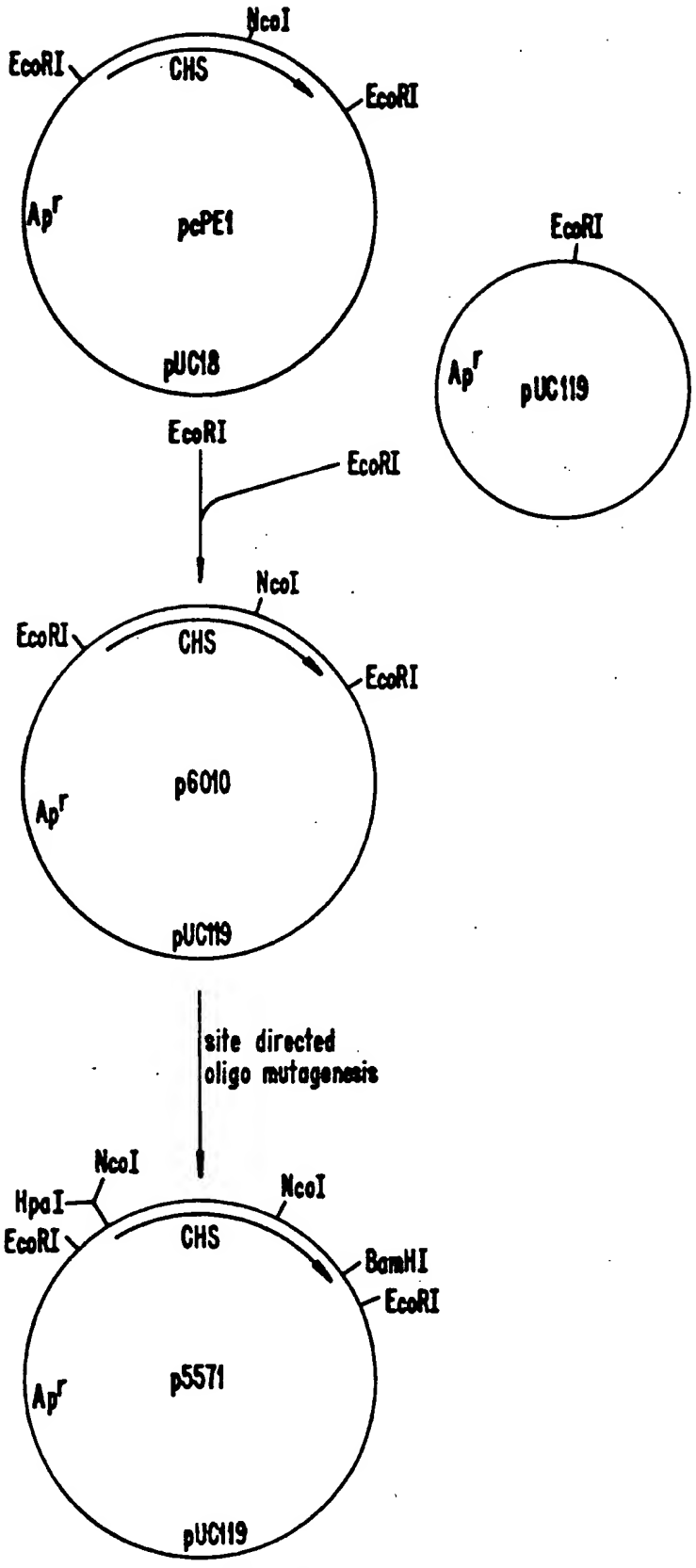

NcoI
EcoRI
CHS
EcoRI
Ap^r
pcPE1
pUC18
EcoRI
Ap^r
pUC119
EcoRI
EcoRI
NcoI
EcoRI
CHS
EcoRI
Ap^r
p6010
pUC119
site directed
oligo mutagenesis
NcoI
HpaI
NcoI
EcoRI
CHS
BamHI
EcoRI
Ap^r
p5571
pUC119

FIG._1A.

GENETIC ENGINEERING OF NOVEL PLANT PHENOTYPES

FIELD OF THE INVENTION

This invention relates generally t the use of recombinant DNA methods for genetically altering plants, and more particularly, to improved means for altering color patterns and color intensity of flowers and other plant parts.

BACKGROUND OF THE INVENTION

The water soluble pigment flavonoids are significant in their contribution to the coloration and other properties of higher plants. For example, the flavonoids are responsible for most orange, scarlet, crimson, mauve, violet and blue colors, and contribute significantly to yellow, ivory and cream colored flowers. See Harborne, (1976) *Chemistry and Biochemistry of Plant Pigments*, 2d ed., Goodwin (Ed.) Acad. Press, London. The most important of the pigment molecules are the anthocyanins, particularly pelargonidin, cyanidin and delphinidin. These are the darker colored pigments responsible for the orange-red, magenta and mauve colors, respectively. The other major flavonoid types, the chalcones, isomeric flavanones, flavones and flavonols are light colored and tend to have relatively smaller effects on intensity or patterns of color.

The functions of these pigments extend well beyond coloration of flowers, however. The pigments also color fruits, leaves and other plant parts, and importantly provide plants with UV protection, as well as protection against herbivores and microbes. Other uses include allelopathy and even some pharmaceutical applications.

The biosynthetic pathways of these various pigments have been extensively studied in many different plant species. The chalcones and aurones are products requiring only the initial biosynthetic enzymes, being direct products of the earliest precursors. The flavones and flavonols are intermediate, and the anthocyanins are products requiring substantial modifications from the initial precursors. All of these products are dependent upon the activity of the initial enzyme chalcone synthase (CHS), which catalyses the production of chalcone from three molecules of malonyl-Coenzyme A and one molecule of coumaroyl-Coenzyme A.

Essentially, all of these phenotypic traits have naturally evolved coordinately with constraints related to plant reproduction. For example, the appearance of a flower has generally resulted from the requirement to attract insects who assist in the pollination process essential for the sexual reproduction of the higher plants. Of course, the decorative and ornamental features impart to flowers a significant commercial value.

Mankind has traditionally intervened in some of the natural processes by, e.g., simply selecting particular flower colors and patterns which might otherwise not have survived in nature. Breeders routinely generate new and unusual flower phenotypes by a variety of time-tested breeding methods. The classical techniques for breeding improved plants, such as different flower varieties with altered flower color intensities or color patterns, typically required natural genetic variability within the experimental gene pool of the species and its relatives. More recently, the generation of variability by induction of mutations has been utilized. Breeders then select among the resulting population those products exhibiting interesting phenotypes, for further characterization.

Unfortunately, the induction of mutations to generate diversity often involves chemical mutagenesis, radiation mutagenesis, tissue culture techniques, or mutagenic genetic stocks. These methods provide means for increasing genetic variability in the desired genes, but frequently produce deleterious mutations in many other genes. These other traits may be removed, in some instances, by further genetic manipulation (e.g., backcrossing), but such work is generally both expensive and time consuming. For example, in the flower business, the properties of stem strength and length, disease resistance and maintaining quality are important, but often initially compromised in the mutagenesis process.

The advent of recombinant DNA technology has provided horticulturists with additional means of modifying plant genome. While certainly practical in some areas, to date genetic engineering methods have had limited success in modifying the flavonoid biosynthetic pathway. Recently, the inhibition of flower pigmentation with a constitutively expressed "anti-sense" chalcone synthase gene has been reported (Van der Krol et al., (1988) Nature 333:866–869).

Thus, there exists a need for improved methods for producing plants with desired phenotypic traits. In particular, these methods should provide general means for phenotypic modification, and may lessen or eliminate entirely the necessity for performing expensive and time-consuming backcrossing.

SUMMARY OF THE INVENTION

In accordance with the present invention, methods are provided for producing plants exhibiting one or more desired phenotypic traits. The invention is based in-part on the surprising discovery that plants exhibiting the desired trait(s) can be selected from transgenotes comprising a DNA segment operably linked to a promoter, wherein transcription products of the segment are substantially homologous to corresponding transcripts of endogenous flavonoid biosynthetic pathway genes. The transgenotes are grown into plants, such as flowering plants capable of exhibiting novel traits, including a reduction in color intensity, an altered pattern color, or a change in basic color of the plant flowers or other plant organs.

The invention further embraces the introduction of one or more flavonoid biosynthetic pathway gene regions, under the control of promoter regions, into dicots and other plants in which the gene is endogenous. In particular, the invention comprises plants, such as those of the genera Petunia and Chrysanthemum, wherein the plant is grown from a cell transformed with a sequence which is transcribed into an RNA sequence substantially homologous (other than anti-sense) to a desired flavonoid biosynthetic pathway gene. DNA or RNA equivalents are introduced into plants in a way to produce more of the endogenous (already present) transcript, but not to produce solely an anti-sense transcript. This is accomplished by using a DNA segment (natural or constructed) in which the promoter is positioned in the normal orientation at the 5' end of the encoding region so that a "sense" transcript (rather than antisense transcript) can be produced. The plant cells can be transformed with a variety of vectors, such as viral vectors, episomal vectors, shuttle vectors, Ti plasmid vectors and the like.

The invention also embraces methods for reducing expression of endogenous nucleic acid sequences coding for proteins acting in a flavonoid biosynthetic pathway of a plant, the method comprising the step of introducing into a cell of the plant a DNA sequence substantially homologous to the endogenous sequence and under the operational control of a promoter sequence, such as a cauliflower mosaic virus sequence. The DNA segment typically comprises at least about 50 nucleotides and may be a full length gene, such as a chalcone synthase coding sequence. In addition, the invention comprises the methods of preparing and using the various DNA constructs of the present invention.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 shows the construction of exemplary plasmids of the present invention.

DETAILED DESCRIPTION

The present invention provides novel methods for producing plants, and embraces the plants so produced, and methods of their use. The invention is based in part on the discovery that a reduction in expression (i.e., repression) of a cellular gene product may be attained upon introduction into the cell of a flavonoid biosynthetic pathway nucleic acid fragment that is ultimately transcribed to yield a mRNA transcript substantially homologous to a portion of the gene's transcript. The duplicated transcript is preferably produced prior to the native transcript, but may be produced simultaneously with the native transcript as well. Depending on the time and amount of transcript produced in a transgenote, a plant grown from it will exhibit a variety of different phenotypic traits. In particular, selecting plants with varying color patterns and intensity, typically without harming other desirable plant characteristics, can be readily achieved in accordance with the present invention.

By way of example, and not limitation, an exemplary preferred embodiment of the present invention entails introducing a full-length chalcone synthase (CHS) coding sequence operably linked to a cauliflower mosaic virus promoter into *Petunia hybrida* cells. These transgenotes are grown into plants and variations in flower coloration are selected. The modified flowers exhibit substantially all of the characteristics of the native *Petunia hybrida* plants. Those skilled in the art will readily appreciate that other traits, other plant nucleic acid sequences and the like may be readily substituted in accordance with the following guidelines.

Traits

A variety of traits are selectable with appropriate procedures and sufficient numbers of transgenotes. Such traits include, but are not limited to, visible traits, environmental or stress related traits, disease related traits, and ripening traits. Among the easiest to select are the flavonoid genes, giving rise to visible traits. In particular, the traits of color intensity, color hue and color pattern are subject to the repression effect.

The class of genes within the flavonoid biosynthetic pathway includes those nucleic acid sequences directly involved in reactions or control of reactions which synthesize or modify a flavonoid compound. Flavonoids are a class of compounds, numbering about 3000 whose functions in plants include coloration in flowers, fruits, leaves, and other organs. Examples of flavonoid biosynthetic genes include those for chalcone synthases, chalcone isomerases (CHI), flavanone 3-hydroxylases, dihydroflavonol reductases, flavanone 2-hydroxylases, dihydroflavanol 2-hydroxylases, flavonoid 3'-hydroxylases, flavonoid 5'-hydroxylases, flavonoid glycosyltransferases (including glucosyl transferases such as UDPG: flavonoid 3-O-glucosyl transferase and UDPG: flavonol 7-O-glucosyl transferase, and rhamnosyl transferases), flavonoid methyltransferases (such as SAM:anthocyanidin 3-(p-coumaroyl)-rutinoside-5-glucoside 3',5'-O-methyltransferase) and flavonoid acyltransferases. See, Hahlbrock (1981) *Biochemistry of Plants*, Vol.7, Conn (Ed.); Weiring and de Vlaming (1984) "Petunia", K.C. Sink (Ed.), Springer-Verlag, N.Y.

Anthocyanin pigmented flowers have colors throughout the range orange to red to purple to blue. Chalcones and aurones are yellow or orange. Flavones and flavonols are very light yellow, or "cream" colored. Flavanones are colorless. Elimination of anthocyanins and diversion of the pathway to flavone or flavonol production would create cream colored flowers. Shifts from blue to purple or purple to red or red to orange can be engineered by interfering with 3' or 5' hydroxylases of 2-hydroxylases. Interference with 2-hydroxylases can also reduce color intensity of anthocyanin pigmented plants. Interference with CHS would create white flowers and with chalcone isomerase would create yellow flowers. A wide variety of bi-color patterns can be created, the two colors being the color of the target plant before engineering and the color resulting from the expression of the introduced flavonoid gene. Types of patterns include: radial star-like patterns; picotee (white outer edge); white annular center; concentric colored rings; erratic, irregular patterns, e.g., variegated or blotchy. There are many variations on these patterns, some more attractive than others, some with sharp boundaries between colors, some with diffuse boundaries, some with linear boundaries, some with wavy, curved boundaries. Also lighter, solid colors are observed.

Suitable sources for flavonoid gene sequences usable in accordance with the present invention are plants, in particular higher plants, virtually all of which normally possess a flavonoid biosynthetic pathway of some type. Any flavonoid pathway which generates naringenin chalcone or compounds generated from naringenin chalcone which itself is generated from coumaroyl-Coenzyme A and malonyl-Coenzyme A by chalcone synthase will be appropriate.

Introduced Nucleic Acid Sequences

The properties of the nucleic acid sequences are varied, and the preferred embodiments will describe a number of features which the person of skill in the art may recognize as not being absolutely essential, but clearly advantageous. These include isolation methods of the particular sequence to be introduced, certain features of the sequence and certain features of the associated vector, if any. Transcriptional expression of the introduced gene is important, and—without intending to be limited to a particular mechanism—additional production of a transcript relative to the normal expression of the endogenous form of the sequence is likely part of the underlying mechanism, especially prior to attainment of peak levels of endogenous gene expression.

RNA resulting from transcription shall be referred to herein on occasion as "transcript" or "mRNA". Typically, transcript which is processed (e.g., introns removed and 3' end polyadenylated) is referred to as

mRNA ("messenger"). As used herein "homolog us" means c rresponding t (the same as). RNA which is homologous to a gene, is RNA which corresponds to the template sequence (with the normal exception of uracil f r RNA in place of thymidine for DNA). Thus, cellularly produced "homologous RNA", as used herein, is not complementary to the template DNA strand of the gene.

Expression of an endogenous gene, e.g., a gene in the flavonoid pathway, yields varying levels of transcript depending on the type of cell and its developmental stage. During flower development, certain cells, e.g., cells that give rise to petal epidermal tissue, produce or begin to produce a transcript at a level which rises at or subsequent to flower meristem initiation. The transcript level reaches a peak later in flower development and eventually decreases. This rise and fall of transcript level may occur over a series of days, e.g., 7–14 days. The rise may also occur rapidly, e.g., over a period of hours, especially in the event of induction such as by UV or visible light. For example, the transcript level is usually decreasing at the mature flower stage (flower maturation).

The proposed mechanism of repression would require that some transcription of the introduced sequence be produced. While the sequence need not necessarily correspond to the final translated message or a part thereof, there are corresponding forms of the mRNA which are functional in repression, but still contain parts of introns or only nontranslated segments of the primary transcript of the normal endogenous sequence. Thus, the effect may occur where the introduced sequence contains no coding sequence per se, but only intron or untranslated sequences homologous to sequences present in the primary transcript of the endogenous sequence.

The introduced sequence generally will be substantially homologous to the endogenous sequence intended to be repressed, such that the controlling elements recognize that the introduced sequence is present, the interaction results in the repressive effect. This minimal homology will typically be greater than about 65%, but a higher homology might exert a more effective repression of expression of the endogenous sequences. Substantially greater homology, or more than about 80% is preferred, though about 95% to absolute identity would be most preferred. Consequently, the effect should apply to any other proteins within a similar family of genes exhibiting homology or substantial homology. For example, the chalcone synthase protein may be encoded by one or more homologous genes which comprise the chalcone synthase gene family, and repression of one member of the family will typically serve to impose the same repressive effect on others of the family. Similarly, for example, chalcone synthase genes from other plant species may be utilized.

The introduced sequence, needing less than absolute homology, also need not be full length, relative to either the primary transcription product or fully processed mRNA. A higher homology in a shorter than full length sequence compensates for a longer less homologous sequence. Furthermore, the introduced sequence need not have the same intron or exon pattern, and homology of non-coding segments will be equally effective. Normally, a sequence of greater than 50–100 nucleotides should be used, though a sequence of greater than about 200–300 nucleotides would be preferred, and a sequence f greater than 500–1000 nucleotides would be especially preferred. However, an introduced sequence identical to the endogenous genomic sequence will be most preferred.

It should be noted that since a full length coding sequence is unnecessary, it is possible t produce the same effect on multiple proteins using a single transf r-mation by fusing multiple sequences t gether to coordinately repress various different genes. Assuming a sufficient number of introductions are made, the introduced sequence need not be linked to an operative promoter sequence. However, a promoter sequence would be preferred, particularly a partially or fully constitutive promoter. "Operably linked" refers to functional linkage between the affecting sequence (such as a promoter or 3' segments) and the controlled nucleic acid sequence. The same effect would be produced by the introduction of a promoter operably linked to the coding strand of an endogenous sequence. This can be effected by either the introduction of a promoter alone to a site operably linked to the target sequence, or by the reintroduction of a sequence of endogenous origin recombinantly attached to an operably linked promoter (resulting in a chimeric gene).

A heterologous sequence is one that originates from a foreign species, or, if from the same species, is substantially modified from its original form.

In considering the expected temporal stage of expression of the introduced gene, relevant factors include the type of promoter, the temporal pattern of the promoter, and the operation of the promoter in view of its position within the genome. A promoter which is expressed concurrently with or prior to the normal activation of the homologous endogenous sequence is preferred. A constitutive promoter is most preferred, such as the cauliflower mosaic virus promoter. This promoter is constitutive because its operation is relatively independent of the developmental stage of the cell in which it is contained. A regulated promoter, such as ones associated with the ribulose-1,5-bisphosphate carboxylase, the chlorophyll binding proteins or the glycine-rich root protein genes are also suitable. This control may be either temporal with respect to the developmental stage of the cell, or based upon differential expression by different parts or organs of the plant.

As referred to above, the operation of a promoter may vary depending on its location in the genome. Thus, a regulated promoter may operate differently from how it does in its normal location, e.g., it may become fully or partially constitutive.

It is preferred to have the DNA sequence linked to and situated at a distance from the promoter corresponding to the distance at which the promoter is normally most effective so as to ensure sufficient transcriptional activity. This distance should be within about 1000 nucleotides, preferably within about 500 nucleotides and optimally within 300 nucleotides of the translation initiation codon.

At the 3' end of the coding sequence, operably linked segments may also be included. Thus, it would be optimum to have a 3' untranslated region containing the polyadenylation site and any relevant transcription termination sites. A 3' sequence of less than about 1000 nucleotides is sufficient, about 500 preferred and about 300, or the length of the 3' untranslated tail of the endogenous sequence is be optimum.

If the introduced gene is an intact gene from the target plant or other plant species (meaning a complete gene containing coding sequences, intron, promoter,

enhancers and other cis-acting regulatory elements either upstream (5') or downstream (3') of the coding sequences), a fraction of independent transgenotes, depending on the gene, may carry the introduced gene in locations that result in abnormal expression, i.e., expression at abnormal times in development. If the introduced gene is a chimeric gene (meaning that one or more elements, such as a promoter, from another gene has been substituted for a component of the intact gene or added to the intact gene, including coding sequences fused to upstream and downstream sequences necessary or beneficial for expression) and is driven by a constitutive (fully or partially) promoter, then abnormal levels and times of expression will be achieved in a large fraction of transgenotes. If the introduced gene is a chimeric gene and is driven by a developmentally regulated promoter, depending on the promoter, some fraction of transgenotes will show abnormal levels and times of expression of the introduced gene. The strength of the promoter or other cis element can be the same, lower, or higher than the coding sequence's usual promoter. The timing in development can be earlier or the same.

While many of these improvements suggested are not essential, the efficiency of production of useful transgenotes may be significantly affected. Some of the transgenotes may be identical to the parental plants, others may have reduced amounts of colored or colorless flavonoids throughout the petals or other organs of interest. Others may have reduced amounts of flavonoids in certain cells or patches of cells or segments of petals or other organs resulting in regular or irregular patterns. Flowers on the same plant may even have different patterns. The likelihood of obtaining a desirable transgenote will depend upon the number of transgenotes screened and the efficiency of actual transformation and expression of the foreign nucleic acid sequence.

The choice of nucleic acid to exert the described repressive effect is broad. Assuming appropriate selection procedures and sufficient numbers of transgenotes, a wide variety of plant genes could display this effect, particularly genes associated with plant pigmentation. Among the plant pigmentation genes are the flavonoid genes, and most particularly the chalcone synthase gene sequence.

Flavonoid gene sequences may be isolated by standard procedures of hybridization of genomic or cDNA libraries by the methods described in Maniatis et al. (see below). Screening may be by (1) nucleic acid hybridization using homologous genes from other organisms, (2) probes synthetically produced to hybridize to particular sequences coding for known protein sequences, or (3) DNA sequencing and comparison to known sequences.

Flavonoid genes may be enriched in libraries by differential hybridization which requires that the mRNA of the target genes be expressed more abundantly in one tissue than in another. Labelled RNA or cDNA from each tissue is hybridized to replicas of the library and tissue specific clones are identified and isolated. Screening can then be used to identify the target gene among the set of tissue specific genes (Kreuzaler et al., (1983) *Proc. Natl. Acad. Sci. USA* 80:2591–2593).

Antibody screening of expression libraries with antibodies made against homologous proteins can select nucleic acid sequences which would code for homologous functions. Selection of sequences homologous to a known flavonoid biosynthetic pathway protein will enable isolation of other forms or equivalent forms from different sources.

Transposon tagging of a flavonoid gene can assist in the isolation of the relevant gene. Transposon tagging involves a mutation of the target gene. A mutation is isolated in which a transposon has inserted into the target gene and altered the resulting phenotype. Using a probe for the transposon, the mutant gene can be isolated. Then, using the DNA adjacent to the transposon in the isolated mutant gene as a probe, the normal wild type allele of the target gene can be isolated (McLaughlin and Walbot (1987) *Genetics* 117:771–776; Dooner et al., (1985) *Mol. Gen. Genetics* 200:240–246; and Federoff et al., (1984) *Proc. Natl. Acad. Sci. USA* 81:3825–3829).

However, as indicated above, the homology between the inserted gene and the endogenous gene need not be absolutely identical. Foreign homologous genes would also be subject to this same repression phenomenon.

Target Plants

As used herein, "plant" refers to either a whole plant, a plant part, a plant cell, or a group of plant cells. The class of plants which can be used in the method of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants. It includes plants of a variety of ploidy levels, including polyploid, diploid and haploid.

The invention has particular application to plants which express the flavonoid pathway genes. At least some of the flavonoid pathway genes are essentially ubiquitous in higher plants; their products are found in flowers or other plant organs (such as leaves, stems, roots, tubers, bracts, sepals, fruits, vegetables) which are colored. These colors are provided largely by anthocyanin pigments, other flavonoid pigments, copigments, or colorless flavonoids synthesized from chalcone by the plant. See Hahlbrock, supra; Harborne, (1986) *Plant Flavenoids in Biology and Medicine: Biochemical Pharmacological and Structure Activity Relationships;* Harborne, (1976) *Chemistry and Biochemistry of Plant Pigments,* (2d ed.) Vol. 1, Goodwin (Ed.) Acad. Press.

Fruit (e.g., apples, cherries, plums, grapes), vegetable (e.g., eggplant, peppers, kale, lettuce, radishes, cauliflower) or other edible plant part (e.g., potato) colors are also subject to manipulation using these techniques. Flower colors, of course, are commonly very dependent on the activity of the flavonoid pathway genes, and thus are especially sensitive to the absolute and relative levels of expression of the flavonoid biosynthetic pathway genes. Ornamental plants and flowers are valuable commercially, and thus are typical targets of the methods herein described. Creation and selection of new coloration schemes are particularly valuable in the ornamental flower bearing plants such as chrysanthemums, carnations, roses, gerberas, lilies, geraniums, poinsettias and petunias.

Transformation

The transformation of plants in accordance with the invention may be carried out in essentially any of the various ways known to those skilled in the art of plant molecular biology. See, in general, *Methods in Enzymology* Vol. 153 ("Recombinant DNA Part D") 1987, Wu and Grossman Eds., Academic Press, incorporated herein by reference. As used herein, the term transformation means alteration of the genotype of a host plant

by the introduction f a nucleic acid sequence. The nucleic acid sequence need not necessarily originate from a different source, but it will, at some point, have been external to the cell into which it is t be introduced.

In one embodiment, the foreign nucleic acid is mechanically transferred by microinjection directly into plant cells by use of micropipettes. Alternatively, the foreign nucleic acid may be transferred into the plant cell by using polyethylene glycol. This forms a precipitation complex with the genetic material that is taken up by the cell (Paszkowski et al., (1984) *EMBO J.* 3:2717–22).

In another embodiment of this invention, the introduced gene may be introduced into the plant cells by electroporation (Fromm et al., (1985) "Expression of Genes Transferred into Monocot and Dicot Plant Cells by Electroporation," *Proc. Natl Acad. Sci. USA* 82:5824). In this technique, plant protoplasts are electroporated in the presence of plasmids or nucleic acids containing the relevant genetic construct. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and form a plant callus. Selection of the transformed plant cells with the transformed gene can be accomplished using phenotypic markers.

Cauliflower mosaic virus (CaMV) may also be used as a vector for introducing the foreign nucleic acid into plant cells (Hohn et al., (1982) "Molecular Biology of Plant Tumors," Academic Press, New York, pp.549–560; Howell, U.S. Pat. No. 4,407,956). The entire CaMV viral DNA genome is inserted into a parent bacterial plasmid creating a recombinant DNA molecule which can be propagated in bacteria. After cloning, the recombinant plasmid again may be cloned and further modified by introduction of the desired DNA sequence into the unique restriction site of the linker. The modified viral portion of the recombinant plasmid is then excised from the parent bacterial plasmid, and used to inoculate the plant cells or plants.

Another method of introduction of nucleic acid segments is high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface (Klein et al., (1987) *Nature* 327:70–73).

A preferred method of introducing the nucleic acid segments into plant cells is to infect a plant cell, an explant, a meristem or a seed with *Agrobacterium tumefaciens* transformed with the segment. Under appropriate conditions known in the art, the transformed plant cells are grown to form shoots, roots, and develop further into plants. The nucleic acid segments can be introduced into appropriate plant cells, for example, by means of the Ti plasmid of *Agrobacterium tumefaciens*. The Ti plasmid is transmitted to plant cells upon infection by *Agrobacterium tumefaciens*, and is stably integrated into the plant genome (Horsch et al., (1984) "Inheritance of Functional Foreign Genes in Plants," *Science*, 233:496–498; Fraley et al., (1983) *Proc. Natl. Acad. Sci. USA* 80:4803).

Ti plasmids contain two regions essential for the production of transformed cells. One of these, named transfer DNA (T DNA), induces tumor formation. The other, termed virulent region, is essential for the introduction of the T DNA into plants. The transfer DNA region, which transfers to the plant genome, can be increased in size by the insertion of the foreign nucleic acid sequence without its transferring ability being affected. By removing the tumor-causing genes so that they no longer interfere, the modified Ti plasmid can then be used as a vector for the transfer f the gene constructs of the invention into an appropriate plant cell, such being a "disabled Ti vect r".

All plant cells which can be transformed by Agrobacterium and whole plants regenerated from the transformed cells can also be transformed according to the invention so as to produce transformed whole plants which contain the transferred foreign nucleic acid sequence.

There are presently at least three different ways to transform plant cells with Agrobacterium:

(1) co-cultivation of Agrobacterium with cultured isolated protoplasts,

(2) transformation of cells or tissues with Agrobacterium, or

(3) transformation of seeds, apices or meristems with Agrobacterium.

Method (1) requires an established culture system that allows culturing protoplasts and plant regeneration from cultured protoplasts.

Method (2) requires (a) that the plant cells or tissues can be transformed by Agrobacterium and (b) that the transformed cells or tissues can be induced to regenerate into whole plants.

Method (3) requires micropropagation.

In the binary system, to have infection, two plasmids are needed: a T-DNA containing plasmid and a vir plasmid. Any one of a number of T-DNA containing plasmids can be used, the only requirement is that one be able to select independently for each of the two plasmids.

After transformation of the plant cell or plant, those plant cells or plants transformed by the Ti plasmid so that the desired DNA segment is integrated can be selected by an appropriate phenotypic marker. These phenotypic markers include, but are not limited to, antibiotic resistance, herbicide resistance or visual observation. Other phenotypic markers are known in the art and may be used in this invention.

All plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred foreign gene. Some suitable plants include, for example, species from the genera Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Hyoscyamus, Lycopersicon, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Ciohorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Hererocallis, Nemesia, Pelargonium, Panicum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Lolium, Zea, Triticum, Sorghum, and Datura.

It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major cereal crop species, sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Limited knowledge presently exists on whether all of these plants can be transformed by Agrobacterium. Species which are a natural plant host for Agrobacterium may be transformable in vitro. Although monocotyledonous plants, and in particular, cereals and grasses, are not natural hosts to Agrobacterium, work to transform them using Agrobacterium has

also been carried out (Hooykas-Van Slogteren et al., (1984) *Nature* 311:763–764). Additional plant genera that may be transformed by Agrobacterium include Chrysanthemum, Dianthus, Gerbera, Euphorbia. Pelaronium, Ipomoea, Passiflora, Cyclamen, Malus, Prunus, Rosa, Rubus, Populus, Santalum, Allium, Lilium, Narcissus, Ananas, Arachis, Phaseolus and Pisum.

Regeneration

Normally, regeneration will be involved in obtaining a whole plant from the transformation process. The term "transgenote" refers to the immediate product of the transformation process and to resultant whole transgenic plants.

The term "regeneration" as used herein, means growing a whole plant from a plant cell, a group of plant cells, a plant part or a plant piece (e.g. from a protoplast, callus, or tissue part).

Plant regeneration from cultural protoplasts is described in Evans et al., "Protoplasts Isolation and Culture," *Handbook of Plant Cell Cultures* 1:124–176 (MacMillan Publishing Co. New York 1983); M.R. Davey, "Recent Developments in the Culture and Regeneration of Plant Protoplasts," *Protoplasts,* (1983)—Lecture Proceedings, pp.12–29, (Birkhauser, Basal 1983); P.J. Dale, "Protoplast Culture and Plant Regeneration of Cereals and Other Recalcitrant Crops," *Protoplasts* (1983)—Lecture Proceedings, pp. 31–41, (Birkhauser, Basel 1983); and H. Binding, "Regeneration of Plants," *Plant Protoplasts,* pp.21–73, (CRC press, Boca Raton 1985).

Regeneration from protoplasts varies from species to species of plants, but generally a suspension of transformed protoplasts containing copies of the exogenous sequence is first made. In certain species embryo formation can then be induced from the protoplast suspension, to the stage of ripening and germination as natural embryos. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is sometimes advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Shoots and roots normally develop simultaneously. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is fully reproducible and repeatable.

Regeneration also occurs from plant callus, explants, organs or parts. Transformation can be performed in the context of organ or plant part regeneration. See, *Methods in Enzymology, supra:* also *Methods in Enzymology,* Vol. 118; and Klee et al., (1987) *Annual Review of Plant Physiology,* 38:467–486.

In vegetatively propagated crops, the mature transgenic plants are propagated by the taking of cuttings or by tissue culture techniques to produce multiple identical plants for trialling, such as testing for production characteristics. Selection of desirable transgenotes is made and new varieties are obtained thereby, and propagated vegetatively for commercial sale.

In seed propagated crops, the mature transgenic plants are self crossed to produce a homozygous inbred plant. The inbred plant produces seed containing the gene for the newly introduced foreign gene activity level. These seeds can be grown to produce plants that would produce the selected phenotype.

The inbreds according to this invention can be used to develop new hybrids. In this method a selected inbred line is crossed with another inbred line to produce the hybrid.

Parts obtained from the regenerated plant, such as flowers, seeds, leaves, branches, fruit, and the like are covered by the invention, provided that these parts comprise cells which have been so transformed. Progeny and variants, and mutants of the regenerated plants are also included within the scope of this invention, provided that these parts comprise the introduced DNA sequences. Progeny and variants, and mutants of the regenerated plants are also included within the scope of this invention.

Vectors

Selection of an appropriate vector is relatively simple, as the constraints are minimal. The apparent minimal traits of the vector are that the desired nucleic acid sequence be introduced in a relatively intact state. Thus, any vector which will produce a plant carrying the introduced DNA sequence should be sufficient. Also, any vector which will introduce a substantially intact RNA which can ultimately be converted into a stably maintained DNA sequence should be acceptable.

Even a naked piece of DNA would be expected to be able to confer the properties of this invention, though at low efficiency. The decision as to whether to use a vector, or which vector to use, will be guided by the method of transformation selected.

If naked nucleic acid introduction methods are chosen, then the vector need be no more than the minimal nucleic acid sequences necessary to confer the desired traits, without the need for additional other sequences. Thus, the possible vectors include the Ti plasmid vectors, shuttle vectors designed merely to maximally yield high numbers of copies, episomal vectors containing minimal sequences necessary for ultimate replication once transformation has occurred, and viral vectors, including the possibility of RNA forms of the gene sequences. The selection of vectors and methods to construct them are commonly known to persons of ordinary skill in the art and are described in general technical references (Methods in Enzymology, supra).

However, any additional attached vector sequences which will confer resistance to degradation of the nucleic acid fragment to be introduced, which assists in the process of genomic integration or provides a means to easily select for those cells or plants which are actually, in fact, transformed are advantageous and greatly decrease the difficulty of selecting useable transgenotes.

Selection

Selection of transgenotes for further study will typically be based upon a visual assay, but may involve biochemical assays of either enzyme activity or product quantitation. Transgenotes will be grown into plants bearing the plant part of interest and the flavonoid gene activities will be monitored by visual appearance or biochemical assays (Northern blots, see, Maniatis (below); Western blots, see, Ausubel (below); enzyme assays and flavonoid compound assays, including spectroscopy, see, Harborne et al., (Eds.), (1975) *The Flavonoids,* Vols. 1 and 2, [Acad. Press]). Appropriate plants will be selected and further evaluated.

The following experimental section is offered by way of example and not by limitation.

EXPERIMENTAL

In general, preparation f plasmid DNA, restriction enzyme digestion, agar se gel electrophoresis of DNA, Southern blots, DNA ligation and bacterial transformation were carried ut using standard methods. (Maniatis et al., *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Laboratory (1982), referred to herein as "Maniatis" and hereby incorporated by reference). Western blots and other standard molecular biology techniques are also described in Ausubel et al., (1987) *Current Protocols in Molecular Biology*, Vols. 1 and 2, and hereby incorporated by reference.

EXAMPLE 1

Plant Transformation Procedures

In the following examples, reagent materials are commercially available, unless otherwise specified. Enzymes used in the cloning procedures are available from commercial sources. All restriction endonuclease reactions are carried out according to manufacturer instructions. Unless otherwise specified, the reaction conditions for other reactions are standard conditions used in the art, as described, for example, in Maniatis. Luria (L) agar and Minimal A (MinA) agar and broth are described in J. H. Miller, *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, New York (1972) (referred to herein as "Miller" and hereby incorporated by reference). Transformations of competent *Escherichia coli* strain DH-1 were performed according to Maniatis. Plasmid DNA was prepared by alkaline extraction according to Maniatis (referred to herein as "mini-prep DNA" or "mini-prep technique"). Site specific oligonucleotide mutagenesis was carried out as described in Geisselsoder, et al., (1987) *BioTechniques* 5:(8), 786–791, except *E. coli* strain BW313 (dut,unq) was used to produce uracil-containing single stranded DNA, in vitro synthesized double stranded DNA was transformed into *E. coli* strain DH-1, and the Klenow fragment of DNA polymerase I was used for the second strand synthesis which was incubated overnight at room temperature.

Antibiotics are abbreviated as follows: Ap is ampicillin, Km is kanamycin, Rif is rifampicin and Tc is tetracycline. Micrograms are referred to herein as ug and milliliters are referred to as ml. Microliter are referred to as ul.

Creation of Agrobacterium Tumefaciens LBA4404/p5972 and p7506

The plasmid pcPE1 was obtained from H. Reif, Max Planck Institut, Koln. This plasmid contained a nearly full length Petunia hybrida chalcone synthase cDNA clone as an EcoRI fragment. The construction strategy that led to the construction of the binary vectors p5972 and p7506 which were used for the reintroduction and expression of this chalcone synthase gene into target plants is shown in FIG. 1. The plasmids shown in this strategy are labeled with only the relevant restriction sites used either in the construction procedure or discussed in the text. The plasmid numbers in the middle of the circles in the figure are the actual number designations given to the plasmids in the construction strategy. The plasmid number in the lower bottom of the circles refers to the cloning vector that gave rise to the relevant clones. For example the first plasmid listed in the strategy is pcPE1. This clone resulted from a ligation of an EcoRI fragment into the commercially available cloning vector pUC18, so pcPE1 is listed in the middle of the circle and pUC18 is listed in the bottom. The restriction enzymes listed by the drawn circles indicate which enzymes were used t digest the plasmids and an arrow indicates that a ligati n reaction took place. Antibiotic resistant genes that were used to select the clones are indicated inside the circles.

The EcoRI fragment containing the complete coding sequence for chalcone synthase protein was recloned into the EcoRI site of plasmid pUC119 (Viera and Messing, *In Methods in Enzymology*. 153(2) eds. Wu and Grossman, San Diego, pp.3–11, 1987) by digesting both plasmids with EcoRI, ligation, and transformation into competent *E. coli* strain DH-1. A plasmid which contained pUC119 and the EcoRI chalcone synthase fragment was identified by restriction mapping and was designated plasmid p6010. Plasmid p6010 was transformed into competent *E. coli* strain BW313 and single stranded DNA containing uracil was isolated (Viera and Messing, ibid.). A 26 base synthetic primer composed of the sequence, 5'-CTTTTTTCTAGTTAACCATGGTGACT-3', and a 24 base synthetic primer composed of the sequence, 5'-CTACTTAGTGGATCCGGCTTATAT-3', were synthesized on an Applied Biosystems 381A DNA synthesizer using the solid phase phosphotriester method. The 26 base primer was used to introduce two new restriction sites, HpaI and NcoI, at the beginning of the coding sequence. The NcoI site overlapped the ATG start codon of the chalcone synthase and would be used for promoter fusions later in the construction strategy. The 24 base primer was used to introduce a BamHI site that overlapped the TAG translation stop codon and would be used later in the construction strategy to fuse the chalcone synthase gene to a poly-adenylation signal sequence. In vitro synthesized double stranded DNA (dsDNA) was synthesized using the two primers and transformed into competent *E. coli* strain DH-1. Ampicillin resistant colonies were screened using mini-prep DNA for new HpaI and NcoI restriction sites that mapped at the beginning of the gene and a BamHI site that mapped at the end of the gene. The plasmid that fulfilled this and further mapping criteria was designated as plasmid p5571.

The next step in the construction strategy served to fuse a 35S Cauliflower Mosaic Virus (herein called CaMV 35S promoter in the text and p35S in the figure) to the beginning of the chalcone synthase coding sequence and a poly-adenylation signal sequence to the end of the coding sequence. Plasmid pJJ2104 is described in Harpster et al., (1988) *Mol. Gen. Genet.* 212:182-190, and was used as the source of the CaMV 35S promoter and poly-adenylation signal sequence. This plasmid has a modified CaMV 35S promoter contained within a BglII and NcoI fragment. The CaMV 35S promoter within plasmid pJJ2104 is fused to the untranslated leader sequence of the photosynthetic 22L chlorophyll a/b binding protein (here in called "Cab22L") to increase transcriptional efficiency (see above reference). The polyadenylation signal sequence is from the nopaline synthase gene (Depicker et al., (1982) *Mol. Appl. Genet.* 1(6):561–573), and is contained within a BamHI and HindIII fragment in the plasmid pJJ2104.

Plasmid p5571 DNA was isolated and cleaved to completion with BamHI and then cleaved with NcoI under conditions to give a partial, incomplete digestion of the DNA because a second NcoI site lies within the

chalcone synthase coding sequence. The DNA was subjected to electrophoresis through 0.5% low melt agarose in standard Tris-Acetate EDTA buffer (described in Maniatis) with ethidium bromide at 0.5 ug/ml incorporated into the agarose. The gel was examined briefly under medium length ultraviolet light using a transilluminator (wave length 312 nm) and a band corresponding to the length (approximately 1200 base pairs) of the chalcone synthase coding sequence was excised from the gel. The gel fragment was weighed to determine the volume and brought to 0.3 M sodium acetate. The equilibrated agarose was heated to 65 degrees for the 10 minutes and then extracted with an equal volume of phenol saturated with 0.1 M Tris-HCl, pH 8. The aqueous phase was removed and extracted twice wit a chloroform-isoamyl alcohol (24:1) mixture and the DNA was then precipitated from the aqueous solution. (All the above techniques were used according to standard conditions as described in Maniatis.) This eluted fragment was combined with plasmid pJJ2104 which was cleaved to completion with NcoI and BamHI and a ligation reaction was set up and incubated for one hour at room temperature and the reaction products subsequently transformed into competent *E. coli* strain DH-1. Plasmid DNA was screened with restriction enzymes to identify the appropriate plasmid containing the 35S CaMV promoter, the chalcone synthase coding sequence, and the poly-adenylation signal sequence. The plasmid p5634 was identified and subjected to restriction digestion to confirm that the plasmid was the correct one.

Two different binary vectors were used in the construction strategy. Plasmids pJJ3942 and pAGS502 are both based on the broad host range cloning vector pRK290 (Ditta et al., (1980) *Proc. Natl. Acad. Sci. USA* 77:7347-7451) and contain a Neomycin Phosphotransferase II coding sequence fused at the 5' end to a nopaline synthase promoter and at the 3' end to an octopine synthase poly-adenylation signal sequence between the left and right TDNA borders (van den Elzen et al., (1985) *Plant Mol. Biol.* 5:141-154). Plasmid pAGS502 contains a polylinker with cloning sites for BamHI, XbaI, HindIII, XhoI, EcoRI and HpaI for insertion of fragments near the TDNA right border. Plasmid pJJ3942 contains HindIII, BamHI, and HpaI as unique cloning sites near the right border. An enhancer-like sequence from the 35S CaMV promoter is contained between the BamHI and HpaI sites. This fragment spans the sequences between positions −45 and −200 of the 35S CaMV promoter to give approximately 200 bases of sequence upstream from the TATAA box.

The entire 35S CaMV promoter, the chalcone synthase coding sequence and the nos poly-adenylation signal sequence are contained within a BglII and HindIII fragment in plasmid p5634. Plasmid p5634 was digested to completion with BglII and HindIII. Two different binary vectors, pJJ3942 and pAGS502, were digested to completion with BamHI and HindIII and each was used in separate ligation reactions with plasmid p5634 digested with BglII and HindIII. The 5'-overhangs generated by the enzymes BglII and BamHI can ligated together but not recleaved by either enzyme. The ligation reactions were transformed into competent *E. coli* strain DH-1 and tetracycline resistant colonies were isolated. DNA was isolated using the mini-prep technique and screened with the appropriate restriction enzymes to isolate pJJ3942 and pAGS502 derivative plasmids that accepted the BglII and HindIII fragment. Further restriction digestions were performed to confirm the identity of the resulting plasmids. The ligation product of pJJ3942 containing the insert was designated as plasmid p5972 and the ligation product of pAGS502 containing the insert was named plasmid p7506.

Plasmids p5972 and p7506 were mobilized (transferred) separately to *Agrobacterium tumefaciens* strain LBA4404 (Ooms et al., (1981) *Gene* 14:33-50). A triparental mating procedure using *E. coli* strain DH-1 carrying p5972 or p7506 (both tetracycline resistant), *E. coli* HB101 carrying plasmid pRK2013 (kanamycin resistant) (Ditta et al., (1980) *Proc. Natl. Acad. Sci. USA* 77:7347-735I) and *A. tumefaciens* strain LBA4404 (rifampicin resistant) was set up. The two *E. coli* strains were grown up overnight on L agar (see Miller) containing the appropriate antibiotics. The *A. tumefaciens* was grown up overnight in MinA broth (see Miller) with no selection. One ml of the *A. tumefaciens* culture was pipetted into a sterile microcentrifuge tube and spun in a microcentrifuge for 2 minutes to pellet the cells. The supernatant was removed and 100 ul of fresh MinA broth was added to resuspend the pellet. A small amount of the *E. coli* cells from each of the overnight cultures was scraped off the petri dish and spread together onto a fresh L agar plate (no antibiotics). The amount of area covered by the cells was approximately 2 cm square. Each amount of *E. coli* cells was approximately equal to the amount of *A. tumefaciens* cells that was collected from 1 ml of culture. The 100 ul of resuspended *A. tumefaciens* cells was added on top of the spread *E. coli* cells and mixed to form a conjugation patch. This petri dish was incubated overnight at room temperature.

On the following day approximately one-fourth of the cells was removed from the conjugation patch and these cells were streaked for single colonies using an L agar plate containing 100 ug/ml rifampicin and 1.2 ug/ml tetracycline. The procedure was repeated four times and resulted in all of the conjugation patch streaked onto four separate plates. These plates were incubated in the dark at room temperature until colonies begin to appear (approximately 3-5 days). Isolated colonies were streaked for single colonies on MinA agar plates containing 1.2 ug/ml tetracycline. The plates were incubated for two days at 28.C. A petri dish containing MinA agar supplemented with 1.2 ug/ml tetracycline was divided into eight equal parts of a circle and eight well isolated single colonies were streaked individually onto sections of the petri dish. This plate was grown up overnight at 28° C. Three-fourths of the cells from each of the eight sections were removed from the agar using a sterile toothpick and the DNA isolated from these cells using the mini-prep technique. Each of the DNAs from these eight preparations was transformed individually into competent *E. coli* strain DH-1 and tetracycline resistant colonies were isolated. One colony from each *E. coli* transformation was grown up and the DNA isolated using the mini-prep technique. The DNA was subjected to restriction enzyme analyses to confirm that the DNA was the original binary clone that was transferred to *A. tumefaciens* LBA4404 via the triparental mating.

Plant transformation

Petunia hybrida varieties: Pink Cascade was obtained from Dr. Michael Reid, Dept. of Environmental Horticulture, University of California, Davis; R18 and V26

were obtained from Dr. Anton Gerats, Dept. of Genetics, Free University, Amsterdam. Petunia hybrida plants were grown from surface-sterilized seed on sterile solidified agar medium f 1/10 the concentration f medium MS of Murashige supplemented with 0.5% sucrose. After germination, seedling t ps were excised by cutting in the hypocotyl region and transferred to MS with 3% sucrose. Plants were maintained at 28° C. under "cool white" fluorescent light at 4-5000 lux, 16h/day.

About six weeks after planting (day 0), leaves were excised, cut with a scalpel blade into pieces about 5mm square and inoculated with *A. tumefaciens* that had been grown overnight in MinA medium supplemented to 0.2% glucose (medium described by J.H. Miller (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, New York) and adjusted to 0.1–0.2 A_{550} units. Inoculated leaf pieces were placed on incubation medium [basal MS medium (MS +3% sucrose +B5 vitamins) +75–100uM acetosyringone, 1mg benzyladenine (BA) per liter, and 0.2 mg indoleacetic acid (IAA)/l]for two days in a sterile transfer hood at room temperature (approx 22° C). On day 2, 25–30 ml of liquid basal MS medium + cefotaxime (500 mg/l) was added to the plates. Plates were then swirled at 70–100 rpm for 30–60 min. Leaf pieces were transferred with the upper epidermis facing up on selection medium (basal MS +BA (1mg/l), IAA 0.2 mg/l, vancomycin (100mg/l)). The plates were sealed with parafilm and incubated at 24° C. under moderate light (3000–5000 lux). On day 14, leaf pieces were transferred to fresh selection medium. On day 28, calli were excised from leaf pieces and transferred to fresh selection medium and shoots were excised and transferred to hormoneless medium [basal MS +vancomycin (100mg/l) and kanamycin (100mg/l)]. On day 42 and following, shoots were excised from calli and transferred to hormoneless medium. After shoot elongation, shoots were excised and dipped in naphthalene acetic acid (NAA) (0.1mg/l) for root development. After rooting, plantlets were transplanted to soil and grown in a greenhouse.

The chimeric CHS gene in p5972 and p7506 was introduced into several varieties of Petunia: (1) a hybrid variety called "Pink Cascade", (2) an inbred, R18, and (3) an inbred, V26. (See, examples 2–4)

EXAMPLE 2

Novel Derivatives of Pink Cascade Petunia

Pink Cascade produces solid pink flowers. Leaf explants from the Pink Cascade variety were transformed with p5972. Six whole plants (CS18201 through 18206) were produced. All had novel flowers. CS18201, 18203, and 18206 gave pure white petal limbs, petal tubes and anthers. CS18202 and 18205 gave flowers with a color pattern: pink wedges at the outer margin of and in the center of petal limbs with the rest of the flower pure white (some flowers on this plant were solid pure white, other flowers had this pattern). CS18204 flowers were a light, blotchy pink.

Progeny of the cross V26 ×CS18202 included: 12 plants with the color of V26 ×Pink Cascade and 6 plants with novel color patterns similar t the patterns of CS18202, but with smaller pigmented sectors on lower petals than upper petals in some progeny. Thus, the production of novel color patterns by the introduced gene is heritable, but the pattern itself may vary among progeny (because Pink Cascade is a hybrid variety, the progeny are genetically heterogenous).

EXAMPLE 3

Novel Derivatives of the R18 Variety Transformed with p5972

Variety R18 produces solid, pale pink flowers. Cells from the R18 variety were transformed with p5972 to produce 14 plants. Nine plants produced flowers with the normal light pink color of R18 flowers. The flowers of five plants had novel patterns. One plant gave pink radial stripes on a solid white background. Another gave mainly pure white flowers, but one flower had some pink wedge similar to wedges on CS18202. The third gave occasional white wedges at petal junctions on a solid pink background. The fourth gave a mixture of pure white flowers and pink with white radial striations (a star-like pattern). The last gave white flowers with pink wedges at the outer margins of petals.

EXAMPLE 4

Novel Derivatives of the V26 Variety Transformed with p5972 and p7506

Variety V26 produces solid, deep violet flowers. Cells from the V26 variety were transformed with p5972 to produce 37 plants. Twenty-eight plants produced flowers colored the same as the V26 parent. Seven plants had flowers with novel patterns; two plants produced pure white flowers. Three plants had pigmented wedges at petal margins similar to CS18202. One had mostly pure white flowers, but some flowers had single, small (3mm) white spots. One plant gave flowers having a beautiful "Cossack dancer" pattern, i.e., a modified radial, star-like pattern. Two plants gave flowers with a somewhat irregular, blotchy pattern of white and purple patches; these flowers, though irregular, looked somewhat like the dancer pattern.

Cells from the V26 variety were transformed with p7506 to produce 20 plants. Seventeen plants produced flowers colored the same as the V26 parent. Three plants produced flowers with color patterns. One plant produced flowers with occasional small white spots. One plant produced one flower with a white tube, while the other flowers were similar to the V26 parent. One plant had flowers which exhibited nearly randomly distributed but sharply defined blotches.

Several transgenotes with white or patterned flowers were crossed to V26. The progeny of a white transgenote produced violet and white flowers in approximately a 1:1 ratio, as expected for a single gene. The progeny of plants with patterned flowers were also patterned or sometimes pure white. Not all plants were identical in flower color intensity or pattern. The penetrance of the flower color phenotypes was complete in the progeny populations of some transgenotes and incomplete in others, i.e., the segregation ratio of solid violet to patterned or solid white was significantly greater than 1 (for incomplete penetrance).

ANALYSES ON PETUNIA HYBRIDA TRANSGENIC PLANTS

RNA Analyses: The steady state levels of both endogenous (wild type) and introduced (35S CaMV driven) chalcone synthase messenger RNA were analyzed in petunia petal using RNase protection analyses (protocol is titled *RNA Transcription.* available from

Stratagene, 11099 North Torrey Pines Road, La Jolla, Calif.).

Petals of six different developmental stages were first harvested from non-transformed V26 petunia plants and from ne transgenic V26 which had white flowers (plant #21838). These stages were defined according to total petal length, degree of pigmentation, and morphology:

State	Length	Pigmentation & Morphology
1	15 mm	no pigmentation (veins only)
2	30	slight flush of light purple around veins
3	40	definite pigmentation from outer surface
4	53	deeper pigmentation, fully extended, still closed
5	58	fully pigmented, just starting to open
6	nd*	freshly mature, fully expanded

*nd; not determined

RNA was isolated from the above described developmental stages for both wild type V26 and transgenic plant #21838. One flower of each stage except stage 1 was sufficient tissue for extracting RNA. For stage 1, eight to ten flowers were combined for the procedure. Petal tissue was frozen in liquid nitrogen and ground to a fine powder in liquid nitrogen using a pre-chilled mortar and pestle. The tissue was added to 1 ml phenol saturated with 0.1 M Tris-HCl, pH 7.5 and 4 ml buffer (100 mM NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1% SDS) and the contents mixed. One ml of chloroform: isoamyl alcohol mixture (24:1) was added and the contents mixed again. The aqueous phase was transferred to a clean tube and extracted a second time with fresh chloroform: isoamyl alcohol mixture. The aqueous phase was transferred to a clean tube, an equal volume of 4 M lithium acetate was added and the contents of the tube were placed on ice for three hours. The RNA was pelleted by centrifugation and the supernatant was removed. The pellet was dissolved in sterile water, the solution was brought to 0.3 M sodium acetate and 2.5 volumes of ethanol were added to precipitate the RNA. The RNA pellet was dissolved in 100 ul of sterile water and the concentration of RNA was determined spectrophotometrically.

Five ug of RNA was used for each protection assay. A 160 nucleotide, radiolabelled anti-sense cab22L-CHS RNA was transcribed in vitro, used as a probe in the protection assays and annealed to the petal RNAs, all as described in the Stratagene protocol. After incubation with single strand specific ribonucleases RNase A and RNase T1, two different protected fragments will remain, a 94 nucleotide fragment representing the endogenous CHS mRNA and a 160 nucleotide fragment representing the introduced chalcone synthase transcript.

An autoradiogram of the RNase protection assays for all six stages of wild-type V26 petunia petal RNAs showed that the chalcone synthase protection fragment was most abundant in stage 3 and stage 4. From this experiment it was determined that the endogenous CHS mRNA is present in petals at all developmental stages examined, gradually increasing in abundance up to stage 4 and then declining to almost undetectable levels in the mature petal.

RNase protection assays on transgenic plant #21838 showed that protection fragments for both the endogenous and the introduced chalcone synthase were present. The relative levels of the endogenous chalcone synthase message followed a similar developmental profile as seen for wild-type plants; however, the overall message levels were substantially reduced, as observed by visual inspection, in each stage from the levels seen in the wild-type V26 plant. In contrast to the wild type chalcone synthase message, the chalcone synthase message from the introduced 35S CaMV promoter was present at a fairly constant low level throughout each developmental time point. This result demonstrates that the introduced CHS gene had the effect of vastly depressing the steady state level of endogenous CHS mRNA.

PROTEIN ANALYSES

Antibodies were raised in a rabbit against chalcone synthase by injecting the rabbit with a fusion protein made in *E. coli.* This fusion protein consisted of wild type betagalactosidase gene with the entire coding sequence of chalcone synthase ligated in-frame to the 3' end of betagalactosidase (Ruther and Miller-Hill (1983) *EMBO J.* 2:(10):1791–1794). Immune antiserum from the rabbit was used in Western analyses to evaluate wild-type and transformed petunia petals. Western analysis were carried out according to manufacturer's instructions using the Proto Blot system from Promega Biotec; but similar techniques are described in Ausubel et al., (supra).

Protein extracts were prepared from purple and white segregants (described above). The same developmental stages as described above were used. Petal tissue was frozen in liquid nitrogen and then ground to a fine powder using a mortar and pestle. The frozen powder was transferred to a glass tissue homogenizer and extraction buffer (50 mM sodium phosphate buffer pH 7.0, 1 mM dithiothreotol, 0.1% Triton X-100, 1 mM EDTA) was used to homogenize the tissue. Protease inhibitors were added to final concentrations of 1 mM phenylmethysulfonyl fluoride and 0.2 mM Leupeptin. Cell debris was removed by centrifugation and the protein content of the supernatant was determined using the Bradford assay (Bradford (1976) *Anal. Biochem.* 72:248–254). Seventeen ug of protein for each sample was loaded onto an 8% polyacrylamide SDS gel.

Western analysis of purple flowered progeny showed that CHS protein was present in the petal extract of all developmental stages. The amount of CHS protein appeared to be the same for stages 1 and 2, increased in stage 3, stayed approximately the same for stages 3, 4 and 5 and decreased slightly for stage 6. Western analysis of white flowered progeny showed that, in comparison to the purple flowered progeny, barely detectable CHS protein was seen in stage 1 and appeared to be more reduced in stages 2, 3, 4, 5 and 6. These analyses showed that while CHS protein could be easily detected in protein extracts from purple flowered progeny, in protein extracts of white flowered progeny CHS protein was reduced to levels where it was barely detectable.

TLC ANALYSIS OF FLAVONOIDS

Thin layer chromatography (TLC) was done to compare flavonoid synthesis in white versus purple flowers from progeny of the cross #21838 ×V26. Mitchell petunia flowers were used as a negative control for anthocyanin synthesis and a positive control for flavonoid synthesis. Flowers of three different lengths were

used as follows: Mitchell (33mm, 43mm, 65mm), purple (33mm, 42mm, 55mm) and white (35mm, 49mm, 57mm). The tubes were assayed separately from the limbs. Tissue was added to 1.0 ml of 2 N HCl, allowed to stand for 2 hours at room temperature and then hydrolyzed f r 20 minutes at 100° C. The supernatant was transferred to a clean tube and 200 ul of isoamyl alcohol was added. The samples were vortexed for at least five seconds and the two phases allowed to separate. Samples were spotted onto a cellulose TLC plate in four separate applications with drying between applications. Two identical plates were set up and run in two different solvent systems; acetic acid/36% HCl/water (30:3:10) and isopropanol/2N HCl (1:1). The two systems discriminate between the anthocyanins. The purple progeny flowers produced both anthocyanins and flavonols. The Mitchell flowers produced flavonols and little or no anthocyanins. The white progeny flowers produced little or no anthocyanins and little or no flavonols of the type produced by the normal purple flowers.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference. The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the claims.

We claim:

1. A method for modifying petunia flower color pattern comprising altering flower pigment production by transforming a petunia cell with a disabled Ti plasmid of *Agrobacterium tumefaciens*, wherein said disabled Ti plasmid comprises a substantially full length coding region f a chalcone synthase (CHS) gene segment which is operably linked t a promoter such that sense transcripts are produced, and wherein the segment is substantially homologous to an endogenous CHS gene and when transcribed in the transgenote is capable of effecting the flower color pattern modification.

2. A method for modifying color pattern in a flower of a petunia plant having an endogenous chalcone synthase (CHS) gene, the method comprising,

growing whole plant transgenotes from a cell, said transgenotes transformed with a recombinant nucleic acid sequence comprising a substantially full length coding region of a CHS gene segment operably linked to a promoter, wherein the segment is linked to the promoter such that sense transcripts are produced, the segment being substantially homologous to the endogenous CHS gene and when transcribed in the transgenote being capable of effecting the flower color pattern modification; and

selecting a plant exhibiting the modified flower color pattern.

3. A method of claim 2 wherein the promoter is constitutive.

4. A method of claim 2 wherein the promoter is a cauliflower mosaic virus 35S promoter.

5. A method of claim 2 wherein the promoter is heterologous.

* * * * *

US005942657A

United States Patent [19]

Bird et al.

[11] Patent Number: 5,942,657

[45] Date of Patent: Aug. 24, 1999

[54] **CO-ORDINATED INHIBITION OF PLANT GENE EXPRESSION**

[75] Inventors: **Colin Roger Bird**, Bracknell; **Rupert George Fray**, Lenton; **Donald Grierson**, Shepshed; **Wolfgang Walter Schuch**, Corwthorne; **Graham Barron Seymour**, Worthing; **Gregory Alan Tucker**, Shepshed, all of United Kingdom

[73] Assignee: **Zeneca Limited**, London, United Kingdom

[21] Appl. No.: **08/335,763**

[22] PCT Filed: **May 13, 1993**

[86] PCT No.: **PCT/GB93/00979**

§ 371 Date: **Jan. 9, 1995**

§ 102(e) Date: **Jan. 9, 1995**

[87] PCT Pub. No.: **WO93/23551**

PCT Pub. Date: **Nov. 25, 1993**

[30] **Foreign Application Priority Data**

May 13, 1992 [GB] United Kingdom 9210273

[51] **Int. Cl.**[6] **A01H 5/00**; C12N 5/14; C12N 15/29; C12N 15/52; C12N 15/82

[52] **U.S. Cl.** **800/205**; 435/172.3; 435/320.1; 435/419; 536/23.2; 536/23.6; 800/DIG. 44

[58] **Field of Search** ..... 435/172.3, 240.4, 435/320.1, 419; 800/205, 250, DIG. 44; 536/23.2, 23.6

[56] **References Cited**

U.S. PATENT DOCUMENTS

5,283,184	2/1994	Jorgenson et al.	435/172.3
5,387,757	2/1995	Bridges et al.	800/205

FOREIGN PATENT DOCUMENTS

0455316	11/1991	European Pat. Off. .

OTHER PUBLICATIONS

Walden et al. (Sep. 1995) Tibtech 13: 324–331.

Primary Examiner—Douglas W. Robinson
Assistant Examiner—Amy J. Nelson
Attorney, Agent, or Firm—Pillsbury Madison & Sutro LLP

[57] **ABSTRACT**

Process for the inhibition of two or more target genes which comprises introducing into the plant a single control gene which has distinct DNA regions homologous to each of the target genes and a promoter operative in plants adapted to transcribe from such distinct regions RNA that inhibits expression of each of the target genes. Constructs suitable for use in the process, as well as cells and plants containing such constructs are disclosed. Specific examples relate to the pectinesterse and polygalacturonase genes.

8 Claims, 5 Drawing Sheets

Fig. 1

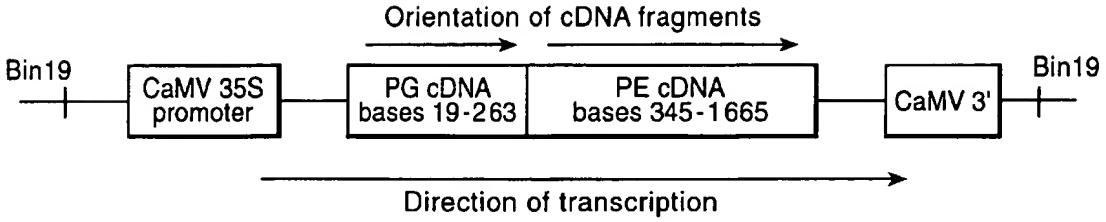

Orientation of cDNA fragments
Bin19
CaMV 35S promoter
PG cDNA bases 19-263
PE cDNA bases 345-1665
CaMV 3'
Bin19
Direction of transcription

Fig. 2A

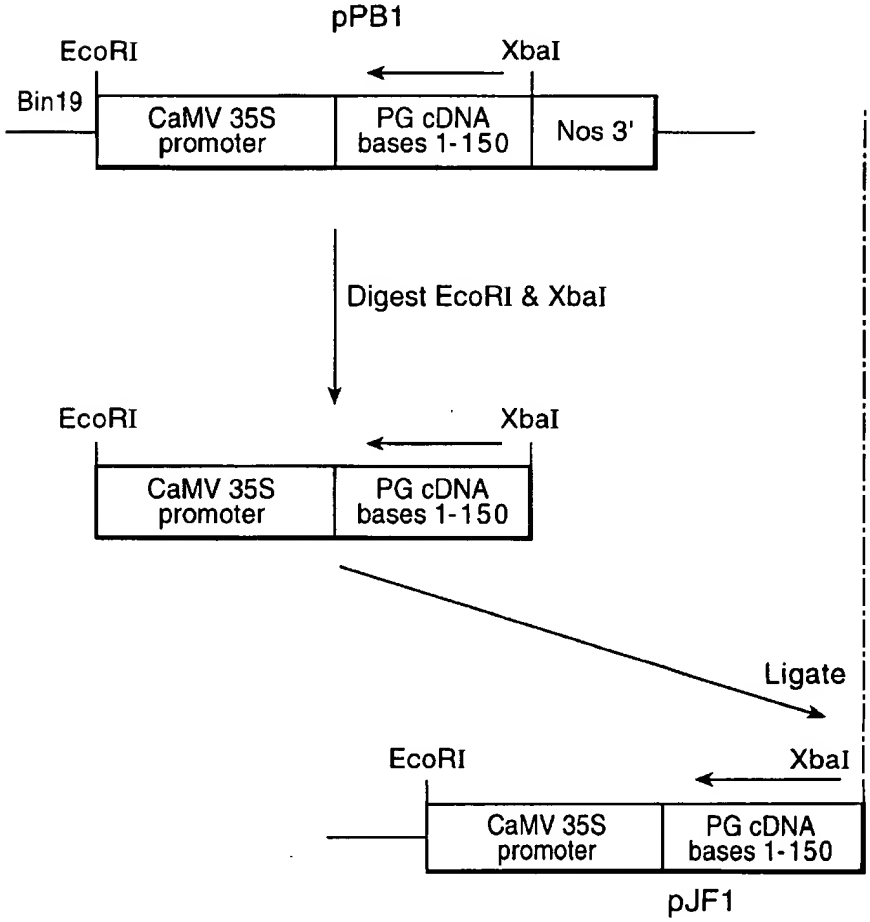

pPB1
EcoRI
XbaI
Bin19
CaMV 35S promoter
PG cDNA bases 1-150
Nos 3'
Digest EcoRI & XbaI
EcoRI
XbaI
CaMV 35S promoter
PG cDNA bases 1-150
Ligate
EcoRI
XbaI
CaMV 35S promoter
PG cDNA bases 1-150
pJF1

Fig. 2B

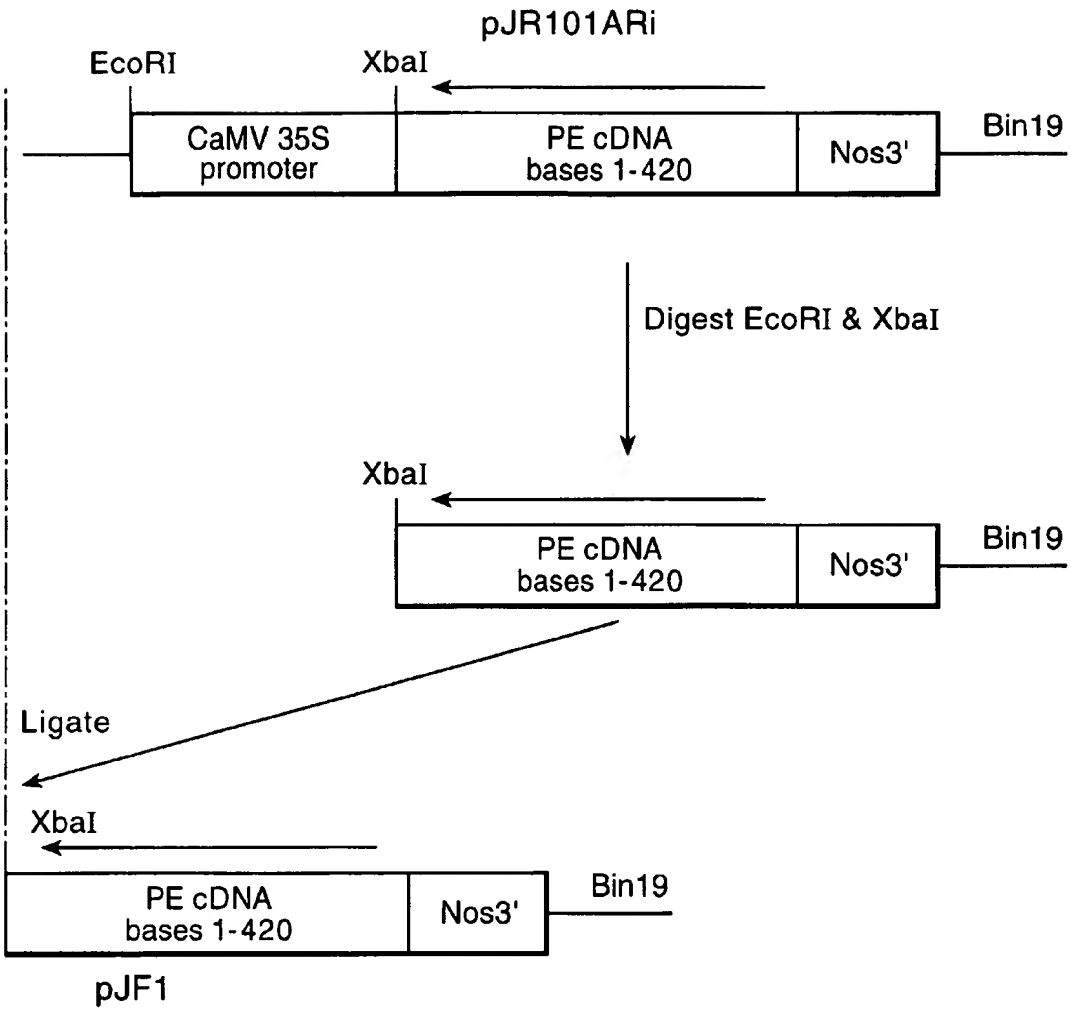

pJR101ARi
EcoRI
XbaI
CaMV 35S promoter
PE cDNA bases 1-420
Nos3'
Bin19
Digest EcoRI & XbaI
XbaI
PE cDNA bases 1-420
Nos3'
Bin19
Ligate
XbaI
PE cDNA bases 1-420
Nos3'
Bin19
pJF1

Fig. 3A

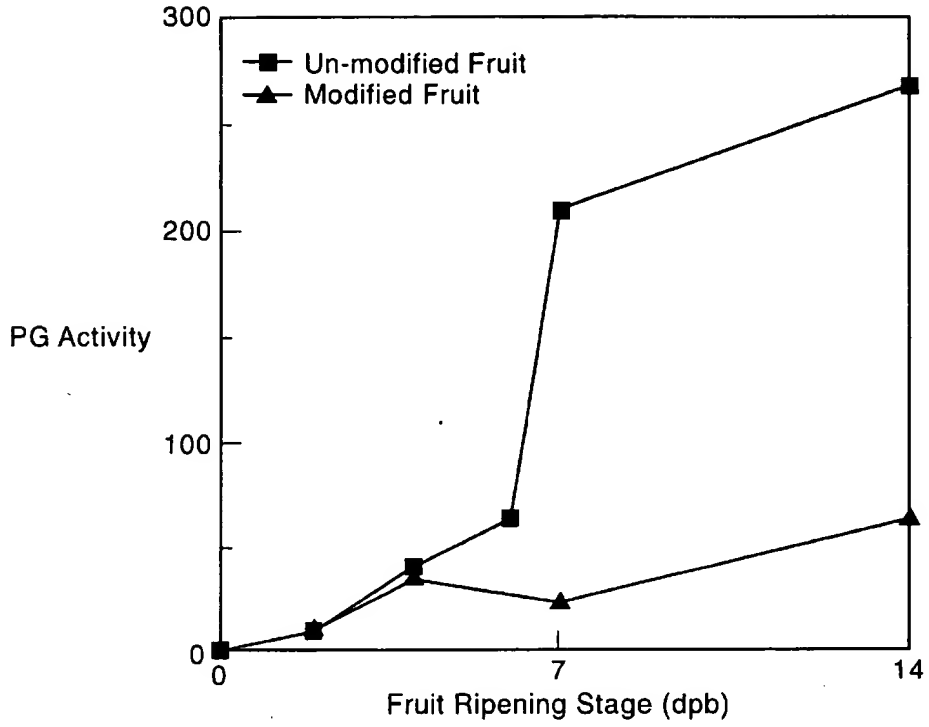
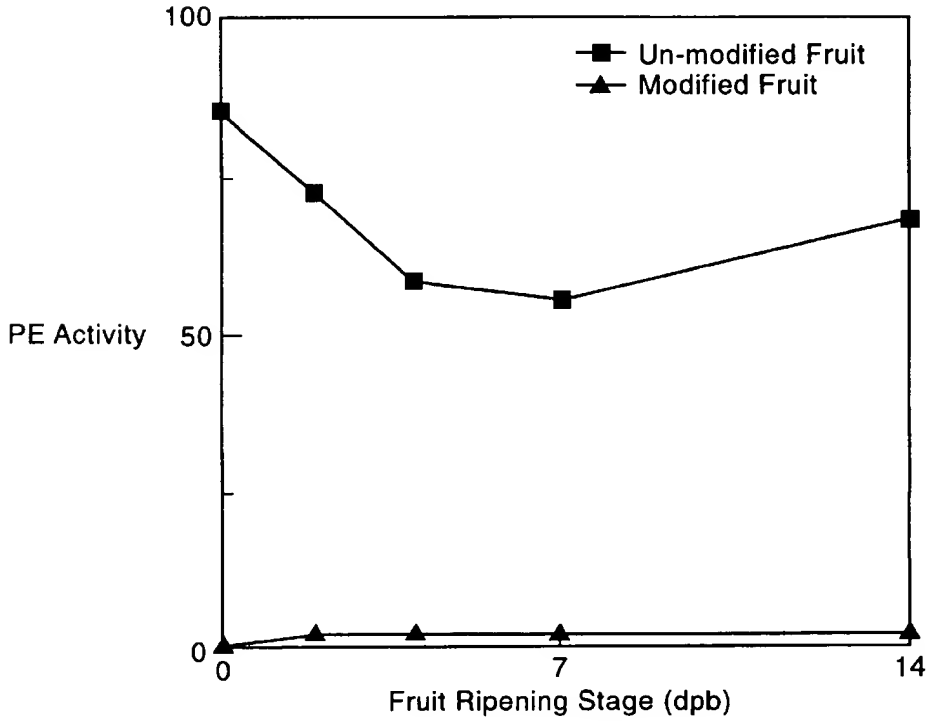

300
200
100
0
PG Activity
Un-modified Fruit
Modified Fruit
0
7
14
Fruit Ripening Stage (dpb)

Fig. 3B


100
50
0
PE Activity
Un-modified Fruit
Modified Fruit
0
7
14
Fruit Ripening Stage (dpb)

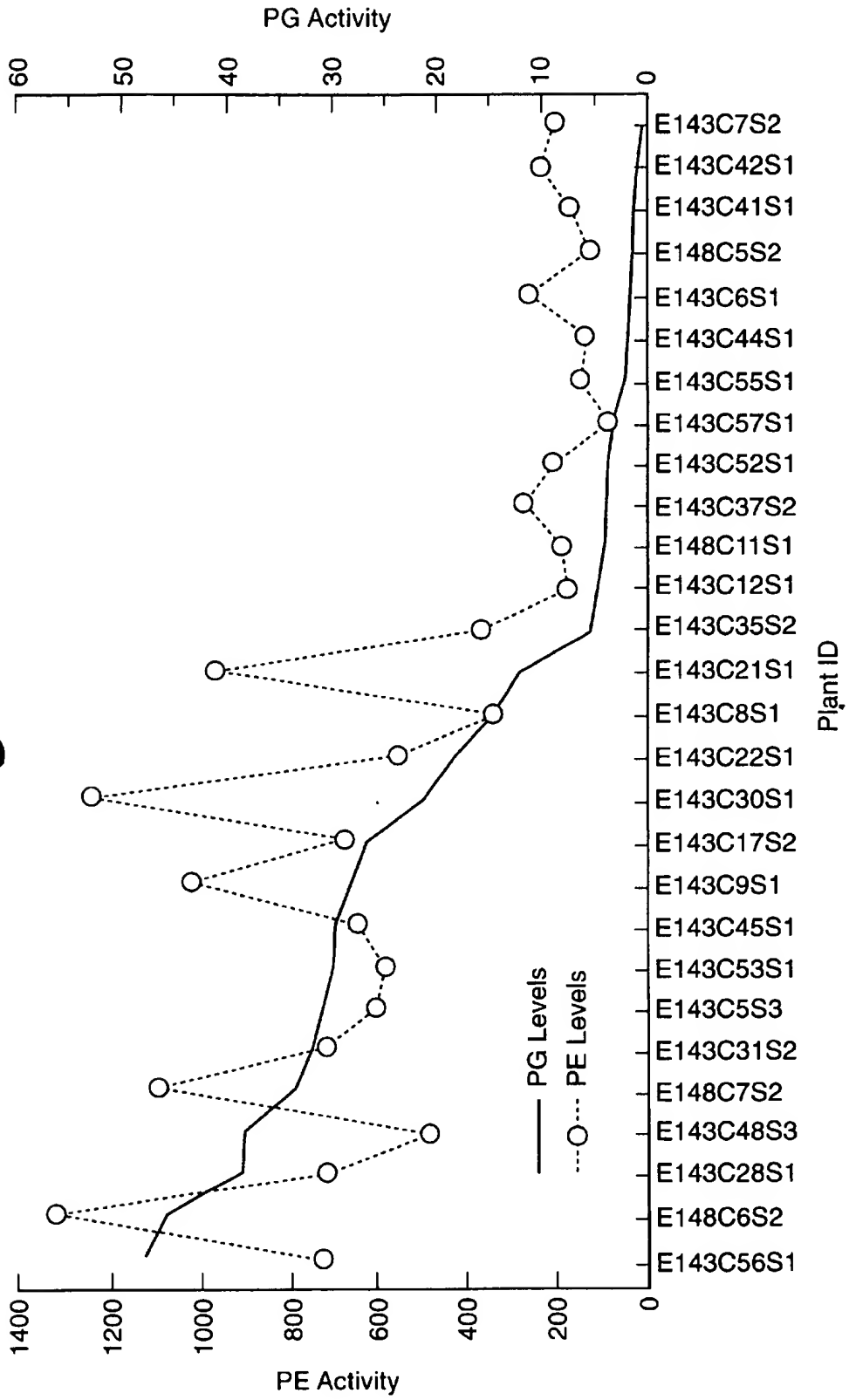

PG Activity
60
50
40
30
20
10
0
E143C7S2
E143C42S1
E143C41S1
E148C5S2
E143C6S1
E143C44S1
E143C55S1
E143C57S1
E143C52S1
E143C37S2
E148C11S1
E143C12S1
E143C35S2
E143C21S1
E143C8S1
E143C22S1
E143C30S1
E143C17S2
E143C9S1
E143C45S1
E143C53S1
E143C5S3
E143C31S2
E148C7S2
E143C48S3
E143C28S1
E148C6S2
E143C56S1
Plant ID
PG Levels
PE Levels
1400
1200
1000
800
600
400
200
0
PE Activity

Fig. 4

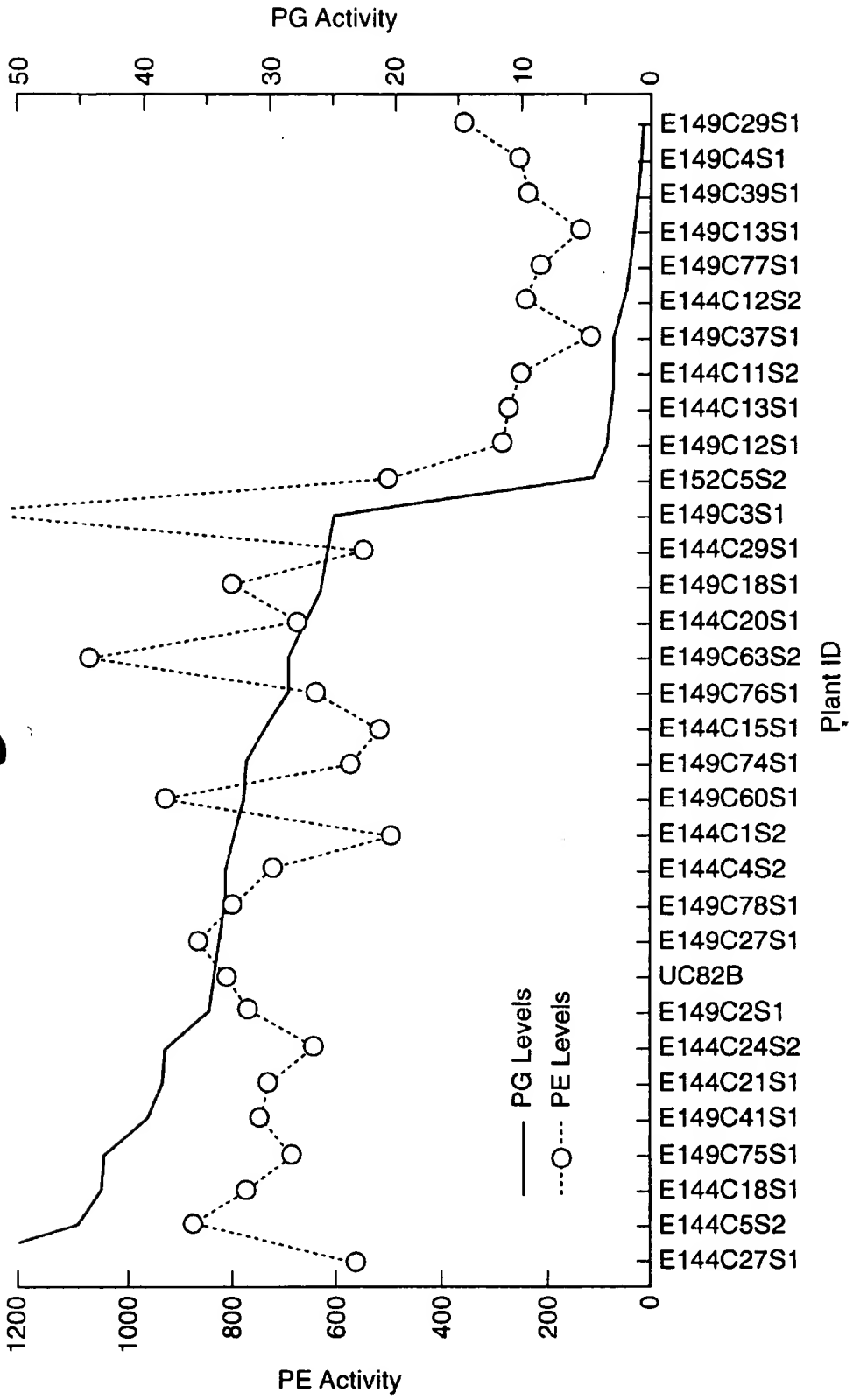

PG Activity
50
40
30
20
10
0
E149C29S1
E149C4S1
E149C39S1
E149C13S1
E149C77S1
E144C12S2
E149C37S1
E144C11S2
E144C13S1
E149C12S1
E152C5S2
E149C3S1
E144C29S1
E149C18S1
E144C20S1
E149C63S2
E149C76S1
E144C15S1
E149C74S1
E149C60S1
E144C1S2
E144C4S2
E149C78S1
E149C27S1
UC82B
E149C2S1
E144C24S2
E144C21S1
E149C41S1
E149C75S1
E144C18S1
E144C5S2
E144C27S1
Plant ID
PG Levels
PE Levels
1200
1000
800
600
400
200
0
PE Activity

Fig. 5

CO-ORDINATED INHIBITION OF PLANT GENE EXPRESSION

This invention relates to novel DNA constructs, plant cells containing the constructs and plants derived therefrom. In particular it involves the use of recombinant DNA technology to control, and more specifically to inhibit, the expression of two or more genes in plants.

BACKGROUND OF THE INVENTION

Plant development is a complex physiological and biochemical process requiring the co-ordinated expression of many genes. The production of new plant varieties with improved agricultural or commercial qualities can be achieved by modifying this coordinated pattern of gene expression. Such modifications have been achieved by conventional plant breeding techniques. However, the exact changes in gene expression that result in the production of the improved variety have not been readily characterised. More recently, recombinant DNA techniques have been used to modify the expression patterns of individual, specific plant genes without directly affecting the expression of other plant genes. In this way, the expression pattern of an individual gene can be either enhanced or inhibited either in the whole plant or in specific tissues or developmental stages.

The inhibition of specific individual plant genes has been achieved by the introduction into the plant of novel genes designed to express RNA homologous, in part, to the endogenous plant gene. In several cases, it has been demonstrated that expression of the target gene can be inhibited by two different strategies. These involve the introduction of specific genes designed to express either antisense or sense RNA. A typical example is the down-regulation of the gene encoding the tomato fruit cell wall enzyme, polygalacturonase, by the expression of either antisense RNA (Smith et al 1988 Nature 344, 724–726) or sense RNA (Smith et al 1990 Mol Gen Genet 224, 477–481). A further example is the down-regulation of the gene encoding chalcone synthase in petunia by either sense or antisense RNA.

The mechanisms by which the expression of a specific gene is inhibited by either antisense or sense RNA genes are not clearly understood. It has been proposed that RNA-RNA duplex molecules may be formed within the cells resulting in the inhibition of expression. However, other and perhaps different mechanisms may operate for the two strategies for down-regulation. Specific individual genes have been inhibited by greater than 99% by the two strategies independently.

SUMMARY OF THE INVENTION

According to the present invention, we provide a process for the inhibition of two or more target genes which comprises introducing into the plant a single control gene which has distinct DNA regions homologous to each of the target genes and a promoter operative in plants adapted to transcribe from such distinct regions RNA that inhibits expression of the target genes. The distinct DNA regions homologous to each of the target genes may be either sense or antisense strands. This invention can be used to generate plants with the combined benefits of down-regulation of several individual genes or families of related genes.

The present invention further provides DNA constructs containing a plant promoter positioned to transcribe an RNA strand from at least two distinct DNA regions homologous to DNA from each of at least two target genes. The distinct DNA regions may be joined sequentially or separated by a spacer region (preferably relatively short) provided such spacer region does not contain a transcription stop signal. The RNA transcribed from this single gene will contain regions homologous to the RNA transcribed from the two or more target genes.

The present invention further comprises novel cells and plants adapted to carry out the process of the invention, or which (or ancestors of which) have been transformed with the constructs of the invention.

The use of DNA constructs according to the invention offers several advantages over alternative, more complex methods of generating plants expressing homologous RNA to two or more individual target genes:

1. Plants can be generated from a single round of transformation thereby eliminating the extra time and complexity of multiple rounds of transformation.

2. The use of a single gene will ensure that RNA homologous to the two or more target genes is transcribed at the same rate. Otherwise, multiple individual homologous genes located in a single construct may have different transcription rates in transformed plants.

3. The segregation patterns of a single, multiple-function gene in progeny of a transformed plant will be simpler than the segregation patterns of multiple, single-function genes that have been combined in the same plant, either by sexual crossing or multiple transformation. This will have significant benefits for subsequent plant breeding.

The genes to which the present invention can be applied include all plant genes for which there is an advantage in down-regulating their activity. In particular the invention can be applied to genes involved in fruit development or ripening-related processes of commercially important fruit-bearing plants, in particular tomato. These could involve combinations of any of the genes encoding: cell wall hydrolases (eg. polygalacturonase, pectin esterase, β-(1-4) glucanase (cellulase), β-galactanase (β-galactosidase)); ethylene biosynthetic enzymes (eg 1-aminocyclopropane-1-carboxylate synthase, ethylene-forming enzyme); carotenoid biosynthetic enzymes (eg prephytoene or phytoene synthase, phytoene desaturase). In addition to genes encoding known enzymes, other genes (eg those showing homology to: pTOM36, pTOM38, pTOM66, pTOM75, pTOM99, pTOM136) with ripening enhanced expression have been identified and may be used in combination with any of the other genes.

DNA constructs according to the invention preferably comprise a base sequence at least 10 bases in length for each of the unrelated target genes for transcription into RNA. There is no theoretical upper limit to the base sequence for each homologous sequence—it may be as long as the relevant mRNA produced by the cell—but for convenience it will generally be found suitable to use sequences between 100 and 1000 bases in length. There is no theoretical limit to the length of the region separating the homologous sequences. The preparation of such constructs is described in more detail below.

The preferred DNA for use in the present invention is DNA derived from CDNA or genomic DNA of the target genes. The required homologous DNA can be obtained in several ways: by cutting with restriction enzymes an appropriate sequence of such DNA; by synthesizing a DNA fragment using synthetic oligonucleotides which are annealed and then ligated together in such a way as to give suitable restriction sites at each end; by using synthetic oligonucleotides in a polymerase chain reaction (PCR) to generate the required fragment with suitable restriction sites at each end. The DNA fragments from each of the unrelated

target genes is then either ligated together prior to cloning or cloned sequentially into a vector containing upstream promoter and downstream terminator sequences.

Recombinant DNA and vectors according to the present invention may be made as follows. Suitable vectors containing the desired base sequences for transcription are treated with restriction enzymes to cut the sequences out. The DNA strands so obtained are cloned either simultaneously or sequentially into a second vector containing the desired promoter sequence (for example cauliflower mosaic virus 35S RNA promoter or the tomato polygalacturonase gene promoter sequence—Bird et al., Plant Molecular Biology, 11, 651–662, 1988) and the desired terminator sequence (for example the 3' of the Agrobacterium tumefaciens nopaline synthase gene, the nos 3' end).

According to the invention we propose to use both constitutive promoters (such as cauliflower mosaic virus 35S RNA) and inducible or developmentally regulated promoters (such as the ripe-fruit-specific polygalacturonase promoter) as circumstances require. Use of a constitutive promoter will tend to affect functions in all parts of the plant: while by using a tissue specific promoter, functions may be controlled more selectively. Thus in applying the invention, e.g. to tomatoes, it may be found convenient to use the promoter of the PG gene (Bird et al, 1988, cited above). Use of this promoter, at least in tomatoes, has the advantage that the production of homologous RNA is under the control of a ripening-specific promoter. Thus the homologous RNA is only produced in the organ in which its action is required. Other ripening-specific promoters that could be used include the E8 promoter (Diekman & Fischer, 1988 cited above).

Vectors according to the invention may be used to transform plants as desired, to make plants according to the invention. Dicotyledonous plants, such as tomato and melon, may be transformed by Agrobacterium Ti plasaid technology, for example as described by Bevan (1984) Nucleic Acid Research, 12, 8711–8721. The process may also be adapted for use with other techniques for generating transgenic plants. Such transformed plants may be reproduced sexually, or by cell or tissue culture.

The degree of production of homologous RNA in the plant cells can be controlled by suitable choice of promoter sequences, or by selecting the number of copies, or the site of integration, of the DNA sequences according to the invention that are introduced into the plant genome. In this way it may be possible to modify expression of the target genes to a greater or lesser extent.

The constructs of our invention may be used to transform cells of both monocotyledonous and dicotyledonous plants in various ways known to the art. In many cases such plant cells (particularly when they are cells of dicotyledonous plants).may be cultured to regenerate whole plants which subsequently reproduce to give successive generations of genetically modified plants. Examples of genetically modified plants according to the present invention include, as well as tomatoes, fruiting plants such as mangoes, peaches, apples, pears, strawberries, bananas and melons.

The specific embodiment of the invention that we have so far studied most thoroughly is the inhibition of both polygalacturonase and pectin esterase in ripening tomatoes. Modified tomato plants have been produced which contain a novel gene designed to express RNA with regions individually homologous to the tomato PG and PE genes. Fruit from some of the primary transformants have significant reductions of both PG and PE activity in the ripening fruit. This reduced enzyme activity is stably inherited as a single gene in progeny of the primary transformants. Plants such as these exhibit the combined benefits of reduced PG and PE activity and will be useful in the production of tomatoes with improved quality for fresh market and processed fruit.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows the structure of the PG-PE sense gene in pPGPE;

FIGS. 2A–2B present a diagram of the strategy for construction of the PG-PE antisense gene in pJF1;

FIGS. 3A–3B are graphs of PG and PE activities in fruit from a tomato plant transformed with pPGPE ("dpb"="days post breaker");

FIG. 4 shows PG and PE activity in fruit from 28 tomato plants transformed with a PE-PG sense construct;

FIG. 5 shows PG and PE activity in fruit from 32 tomato plants transformed with a PE-PG antisense construct.

DETAILED DESCRIPTION OF THE INVENTION

EXAMPLE 1

Construction of Combined PG-PE Sense Gene

A PG-PE sense gene was constructed by cloning bases 19 to 263 of the PG cDNA (pTOM6—Grierson et al 1986 Nuc. Acids Res. 14, 8595–8603) and bases 345 to 1665 of the PE cDNA (pPE1—Ray et al 1988 Eur. J. Biochem 174, 119–124) into the multiple cloning site of the vector pDH51. pDH51 is a pUC based vector which contains a multiple cloning site between the CaMV 35S promoter sequence and the CaMV 3' polyadenylation sequence. The orientation of the PG and PE sequences was determined by restriction mapping and DNA sequence determination of the borders of the inserted fragments. After verification of the structure of the vector, the expression module was isolated by digestion with pvuII and transferred to the plant transformation vector Bin19 (Bevan 1984 Nuc. Acids Res. 12, 8711–8721) to give the vector pPGPE (also known as pSB1).

EXAMPLE 2

Construction of Combined PG-PE Antisense Gene

A PG-PE antisense vector (PJF1) was constructed as shown in FIGS. 2A and 2B. A 686 bp EcoRI-XbaI fragment from pPBI that contained the CaMV 35S promoter sequence and the first 150 bases from the 5' end of the PG cDNA (pTOM6) was isolated. The fragment was cloned into the EcoRI-XbaI sites of the PE antisense vector pJR101ARi after removal of the CaMV 35S promoter sequence. The construction of the vector was confirmed by PCR and DNA sequence analysis.

EXAMPLE 3

Generation of Transformed Tomato Plants

Vectors from Example 1 were transferred to *Agrobacterium tumefaciens* LBA4404 (a micro-organism widely available to plant biotechnologists) and were used to transform tomato plants. Transformation of tomato stem segments followed standard protocols (e.g. Bird et al Plant Molecular Biology 11, 651–662, 1988). Transformed plants were identified by their ability to grow on media containing the antibiotic kanamycin. Plants were regenerated, grown to maturity and the presence of the PG-PE sense gene was confirmed by genomic Southern analysis.

EXAMPLE 4

Analysis of Fruit from Transformed Plants

The PG and PE activities were assayed in ripening fruit from 3 plants transformed with pPGPE and several un-modified plants. The PG and PE activity was reduced by various extents in the fruit from all three transformed plants. The results for fruit from one of the transformed plants are shown in FIGS. 3A and 3B.

EXAMPLE 5

Analysis of Progeny of Transformed Plants

Self-fertilised seed from the pPGPE-transformed plant with the greatest reductions in PG and PE activities were germinated in the presence of kanamycin. Plants that were resistant to kanamycin were grown to maturity and enzyme activities in the fruit were analysed. These plants also had reduced PG and PE activity. Thus, the inheritance of the reduced PG and PE phenotype was confirmed.

EXAMPLE 6

The sense construct pSB1 prepared in Example 1 was used to transform tomatoes of the variety UC82B, using the same protocol as in Example 3. Transformants were grown in the glasshouse. PG and PE enzyme activities were assayed in ripe fruit. Fruit from the first 28 plants to mature were analysed. The results are presented graphically in FIG. **4**, ranked in order of decreasing PG level. These results confirm that plants with substantial reductions in both PG and PE activity were obtained from the population of transformants. In all plants with significantly reduced activity of one enzyme, the other was also reduced. Thus the sense down-regulation of the two genes was tightly linked.

EXAMPLE 7

Expression of Genes in Plants Containing Sense Constructs

Selfed progeny were grown from two primary transformants obtained from Example 3 with single sites of insertion (E148C11-PG expression 10% that of normal fruit, PE expression 20% normal: E143C44-PG expression 5% normal, PE expression 16% normal). Stable inheritance of the double sense gene was confirmed by genomic Southern analysis. Homozygous and azygous progeny were identified. Assays of ripe fruit PG and PE activities confirmed that the low PG and PE phenotype segregated with the sense gene:

		PG activity	PE activity
E143C44.16	Homozygous	1	108
E143C44.30	Azygous	34	307
E148C11.1	Homozygous	2.5	98
E148C11.25	Azygous	34	480
		(all nmol/hr/μg protein)	

EXAMPLE 8

Expression of Genes in Plants Containing Antisense Constructs

68 Transformants with pJF1 (from Example 2) in were generated in the tomato variety UC82B. PG and PE enzyme activities were assayed in ripe fruit. Fruit from the first 32 plants to mature were analysed. Untransformed UC62B is included as a control. The results are presented graphically in FIG. **5**, ranked in order of decreasing PG level. As with the sense plants, reduced activities of the two enzymes were tightly linked.

This data confirms the utility of the antisense construct pJF1 (Example 2).

Analysis of selfed progeny confirms that both the antisense gene and the reduced PG/PE phenotype are inherited.

We claim:

1. A process for the inhibition of polygalacturonase and pectin esterase in a tomato plant which comprise introducing into the plant a single control gene which has distinct DNA regions homologous to the tomato polygalacturonase and tomato pectin esterase genes and a promoter operative in tomato plants adapted to transcribe from such distinct regions RNA that inhibits expression of each of the polygalacturonase and pectin esterase genes and allowing the thus transformed plant to grow, the DNA region homologous to the tomato polygalacturonase gene comprising at least a 100 base sequence of the pTOM6 gene and the DNA region homologous to the tomato pectin esterase gene comprising at least a 100 base sequence of the pPE1 gene.

2. A process as claimed in claim **1** in which at least one of the DNA regions is adapted to produce sense RNA.

3. A process as claimed in claim **1** in which at least one of the DNA regions is adapted to produce antisense RNA.

4. A process as claimed in claim **1** in which the promoter is a constitutive promoter.

5. A process as claimed in claim **1** in which the promoter is an inducible promoter.

6. A DNA construct comprising a single control gene which has distinct DNA regions homologous to the tomato polygalacturonase and tomato pectin esterase genes and a plant promoter positioned to transcribe an RNA strand from said two distinct DNA regions to inhibit expression of each of said genes, the DNA region homologous to the tomato polygalacturonase gene comprising at least a 100 base sequence of the pTOM6 gene and the DNA region homologous to the tomato pectin esterase gene comprising at least a 100 base sequence of the pPE1 gene.

7. A DNA construct as claimed in claim **6** in which at least one of the DNA regions is adapted to produce sense RNA.

8. A DNA construct is claimed in claim **6** in which at least one of the DNA regions is adapted to produce antisense RNA.

* * * * *

Plant Physiology, April 1999, Vol. 119, pp. 1187–1197, www.plantphysiol.org © 1999 American Society of Plant Physiologists

Posttranscriptional Gene Silencing in Transgenic Sugarcane. Dissection of Homology-Dependent Virus Resistance in a Monocot That Has a Complex Polyploid Genome[1]

Ivan L. Ingelbrecht, James E. Irvine, and T. Erik Mirkov*

Department of Plant Pathology and Microbiology (I.L.I., T.E.M.), and Department of Soil and Crop Sciences (J.E.I.), The Texas A&M University Agricultural Experiment Station, 2415 East Highway 83, Weslaco, Texas 78596

RNA-mediated, posttranscriptional gene silencing has been determined as the molecular mechanism underlying transgenic virus resistance in many plant virus-dicot host plant systems. In this paper we show that transgenic virus resistance in sugarcane (*Saccharum* spp. hybrid) is based on posttranscriptional gene silencing. The resistance is derived from an untranslatable form of the sorghum mosaic potyvirus strain SCH coat protein (CP) gene. Transgenic sugarcane plants challenged with sorghum mosaic potyvirus strain SCH had phenotypes that ranged from fully susceptible to completely resistant, and a recovery phenotype was also observed. Clones derived from the same transformation event or obtained after vegetative propagation could display different levels of virus resistance, suggesting the involvement of a quantitative component in the resistance response. Most resistant plants displayed low or undetectable steady-state CP transgene mRNA levels, although nuclear transcription rates were high. Increased DNA methylation was observed in the transcribed region of the CP transgenes in most of these plants. Collectively, these characteristics indicate that an RNA-mediated, homology-dependent mechanism is at the base of the virus resistance. This work extends posttranscriptional gene silencing and homology-dependent virus resistance, so far observed only in dicots, to an agronomically important, polyploid monocot.

Sugarcane (*Saccharum* spp. hybrid) ranks among the world's top 10 food crops and annually provides 60% to 70% of the sugar produced worldwide (Sugar and Sweetener Situation and Outlook Yearbook, 1997). Modern commercial sugarcane cultivars are interspecific hybrids derived from crosses of noble sugarcane, i.e. *Saccharum officinarum* L. ($2n$ = 70–122). Crosses are most often made with *Saccharum spontaneum* L. ($2n$ = 36–128), sometimes with *Saccharum barberi* ($2n$ = 60–140) or *Saccharum sinense* ($2n$ = 104–128), and rarely with *Saccharum robustum* ($2n$ = 66–170) (Irvine, 1999). Sugarcane cultivars have ploidy levels that range from 5× to 14× (× = 5, 6, 8, 10, 12, or 14) and chromosomal mosaicism has been reported (Burner and Legendre, 1994, and refs. therein). The genetic complexity and low fertility of sugarcane render traditional breeding laborious and make it a prime candidate for improvement through genetic engineering.

Transgenic sugarcane plants have been obtained via particle gun bombardment of embryogenic callus (Bower and Birch, 1992; Gallo-Meagher and Irvine, 1996) and via electroporation of cells derived from embryogenic callus (Arencibia et al., 1995). Unlike many other members of the Poaceae in which regeneration is restricted to certain genotypes, most sugarcane cultivars tested to date have yielded regenerable calli. Therefore, introducing specific genetic improvements, such as virus resistance, directly into elite sugarcane varieties is a realistic goal.

SCMV has a monopartite, positive-strand RNA genome (Shukla et al., 1994). Recent taxonomic studies have shown that the SCMV complex comprises four or five different potyviruses, including strains of SrMV, SCMV, Johnson grass mosaic virus, and maize dwarf mosaic virus (Yang and Mirkov, 1997, and refs. therein). Members of the SCMV complex can cause mosaic symptoms and yield loss in susceptible members of sugarcane, maize, sorghum (*Sorghum bicolor*), and other poaceous plants. The worldwide epidemic spread of sugarcane mosaic in the 1920s caused a near collapse of the sugar industry in Louisiana, Argentina, and Brazil and was the main factor in replacing the highly susceptible noble cultivars with interspecific hybrid derivatives that were tolerant to SCMV. However, strains of SCMV and SrMV continue to cause losses in the sugarcane industry.

Several strategies have been used to engineer virus resistance in plants (for review, see Baulcombe, 1996). In CP and movement protein-mediated protection, a transgene-derived homolog of a viral protein is expressed in plants, which interferes with or prevents various stages of the viral life cycle, resulting in attenuated disease symptoms or resistance.

[1] This work was supported by the Texas Higher Education Coordinating Board Advanced Technology Program (grant nos. 999902-029 and 999902-188) and a grant from the International Consortium for Sugarcane Biotechnology.

* Corresponding author; e-mail e-mirkov@tamu.edu; fax 1–956–968–0641.

Abbreviations: CP, coat protein; cRNA, complementary RNA; HDR, homology-dependent resistance; ORF, open reading frame; PTGS, posttranscriptional gene silencing; SCMV, sugarcane mosaic virus; SrMV, sorghum mosaic virus; 3′UTR, 3′-untranslated region.

It is now well established that transgenes in plants can suppress expression of homologous endogenous genes, transgenes, or viral RNAs (for reviews, see Depicker and Van Montagu, 1997; Stam et al., 1997). Homology-dependent gene silencing can occur at the level of transcription or by a posttranscriptional process. Transcriptional silencing is often associated with increased methylation and inactivation of the promoter sequences of the affected genes. In PTGS, an unidentified cellular process is responsible for the specific degradation of the homologous RNA molecules. The expression level, number, and configuration of the integrated transgenes, as well as other less understood developmental and environmental factors, can all influence the occurrence of PTGS. There is now convincing evidence that PTGS is involved in many cases of RNA-mediated virus resistance; this type of resistance has been termed HDR (Mueller et al., 1995).

Recent experiments have shown that a signaling molecule is involved in the systemic spread of gene silencing and resulting virus resistance (Palauqui et al., 1997; Voinnet and Baulcombe, 1997). Gene silencing and virus resistance can be induced by virus infections of nontransgenic plants (Covey et al., 1997; Ratcliff et al., 1997; Al-Kaff et al., 1998). These results and others led to the hypothesis that gene silencing is a natural plant defense mechanism (for review of transgene silencing as a manifestation of cellular defense responses, see Matzke and Matzke, 1998). The recent demonstration that the expression of the P1/HC-Pro sequence of the tobacco etch potyviral genomic RNA can suppress PTGS provides direct evidence that PTGS functions as a host defense response and shows that viruses have evolved a mechanism to counter it (Anandalakshmi et al., 1998; Kasschau and Carrington, 1998).

PTGS and HDR have been conclusively demonstrated for a number of dicot plants with a variety of viruses (for review, see Baulcombe, 1996). Typically, this type of resistance is highly strain specific and two classes of resistance response are often encountered: a pre-established immune phenotype and a delayed recovery phenotype. Resistance is associated with reduced steady-state transgene mRNA levels, while the transgenes remain actively transcribed in the nucleus. Recent data indicate that transgene transcription is required for HDR (Sijen et al., 1996; English et al., 1997; Waterhouse et al., 1998). In a few studies, shorter than full-length transgene-derived RNA transcripts have been detected that may represent degradation products of the silencing mechanism (Goodwin et al., 1996; Tanzer et al., 1997). Finally, some cases of HDR are correlated with methylation in the transcribed region of the silenced transgenes (English et al., 1996; Sijen et al., 1996).

We have transformed sugarcane with an untranslatable form of the SrMV strain SCH CP gene using particle gun bombardment. Phenotypical and molecular data show that in virus-resistant sugarcane plants resistance relies on an RNA-mediated, homology-dependent mechanism. Our studies extend PTGS and its association with virus resistance to a monocot and reinforce the hypothesis of an ancestral PTGS pathway as a defense mechanism in plants.

MATERIALS AND METHODS

Constructs

Plasmid pAHC20 (Christensen and Quail, 1996) contains the *bar* (*bialaphos-resistance* gene)-coding region under the control of the maize ubiquitin promoter, first exon, and first intron, followed by the *nos* (*nopaline synthase* gene) terminator and will be referred to as *Ubi-bar* in this paper. To construct *Ubi-npt*, a 963-bp *Bgl*II/*Apa*LI blunt-ended fragment from pCRII (Invitrogen, Carlsbad, CA) that contains the *npt*-coding sequence was inserted into the blunt-ended *Sal*I sites of *Ubi-bar* to replace the *bar*-coding region. For construction of *Ubi-hut*, which contains the SrMV-SCH untranslatable CP ORF (hut), a 1.1-kb PCR product containing the SrMV-SCH CP-coding sequence and 77 nucleotides of the 3′UTR, was cloned into pCRII using the TA cloning kit (Invitrogen) yielding pSCH-1. Then, the CP ORF and 3′UTR sequences were isolated from pSCH-1 as a 1.1-kb blunt-ended *Eco*RI fragment and inserted into the blunt-ended *Sal*I sites of *Ubi-bar* to replace the *bar*-coding region, yielding *Ubi-hut* (Fig. 1).

The following phagemid clones were constructed for in vitro RNA synthesis and production of single-stranded DNA: (a) a 498-bp *Sph*I fragment from the sugarcane Rubisco small subunit genomic clone *scrbsc-1* (nucleotide position 1313–1811, see Tang and Sun, 1993) was subcloned in a 5′ to 3′ and a 3′ to 5′ orientation into the *Sph*I site of pGem11Zf(+), yielding pIVING366-1 and pIVING366-2, respectively; (b) a 1.1-kb *Eco*RI fragment encompassing the SrMV-SCH CP-coding region and 77 nucleotides of the 3′UTR was isolated from pSCH-1 and subcloned in a 5′ to 3′ and a 3′ to 5′ orientation into the *Eco*RI site of pGem7Zf(−), yielding pIVING362-3 and pIVING362-4, respectively; (c) a 0.7-kb *Pst*I fragment encompassing the 3′ part of the *npt*-coding region was isolated from *Ubi-npt*, rendered blunt with T4 DNA polymerase, and subcloned in a 5′ to 3′ and a 3′ to 5′ orientation into the *Sma*I site of pGem7Zf(−), yielding pIVING362-5 and pIVING362-6, respectively; and (d) a 0.6-kb *Pst*I fragment containing the *bar*-coding region was isolated from *Ubi-bar*, made blunt with T4 DNA polymerase and subcloned in a 5′ to 3′ and a 3′ to 5′ orientation into the *Sma*I site of pGem7Zf(−), yielding pIVING362-7 and pIVING369-1, respectively. These plasmids were transformed into DH5αF′IQ cells (GIBCO-BRL). Unless stated otherwise, all recombinant DNA and bacterial manipulations were performed using standard techniques (Sambrook et al., 1989).

Sugarcane Transformation

Embryogenic callus cultures were established from young leaf bases and immature flowers of the commercial sugarcane (*Saccharum* spp. hybrid) varieties CP65-357 and CP72-1210. Transformation of callus by particle gun bombardment and regeneration of shoots were done as described previously (Gallo-Meagher and Irvine, 1996); 7- to 40-week-old calli were bombarded with an equimolar mixture of either *Ubi-bar* and *Ubi-hut* or *Ubi-npt* and *Ubi-hut* (4 μg DNA/480 μg particles). Two days after bombardment,

calli were transferred to CI-3 medium containing either 1 mg/L bialaphos or 15 mg/L geneticin. Four weeks later, calli were transferred to 2K5N medium containing 2 mg/L kinetin, 5 mg/L NAA, and 1 mg/L bialaphos or 15 mg/L geneticin to promote shoot regeneration and inhibit development of nontransgenic tissue. Tissues were subcultured on this medium every 2 to 3 weeks for approximately 12 weeks. At this time, shoots were placed into Magenta boxes containing RG-2 with 4 mg/L indole butyric acid and either 3 mg/L bialaphos or 45 mg/L geneticin to induce roots. After about 8 weeks, shoots (5–15 cm) with well-developed roots were transferred to peat pots containing potting soil (Metromix, Scotts, Hope, AR) and placed in an environmental growth chamber at 30°C under 15 h of fluorescent and incandescent light. After 2 to 4 weeks, plants (15–30 cm high) were transferred to 15-cm-diameter pots and placed in the greenhouse.

Virus Isolates, Inoculation, and Resistance Tests

The virus strains SCMV-D, SrMV-SCM, SrMV-SCI, and a Texas isolate of SrMV-SCH were previously described (Yang and Mirkov, 1997). Virus strains were propagated on sorghum (*Sorghum bicolor* cv Rio). Sugarcane plants at the four- to eight-leaf stage were inoculated with sap extract (20 mL of a 0.1 M Na_2SO_3/0.1 M phosphate buffer, pH 7.0, containing 0.2 g of diatomaceous earth [Sigma] g^{-1} tissue) from virus-infected sorghum leaves. The plants were inoculated at least twice, at 2- to 3-week intervals. The infection rate on nontransgenic controls was more than 95%. Symptoms were scored visually and, in some cases, the virus level in leaves was quantified by ELISA, according to standard procedures (Converse and Martin, 1990), using SrMV-SCH alkaline-conjugated IgG at a 1:1000 dilution or by RNA gel-blot analysis.

DNA Gel-Blot Analysis

Total DNA was extracted from sugarcane leaves as described by Tai and Tanksley (1990) with slight modifications. Ten micrograms of digested DNA samples was electrophoresed on a 0.8% (w/v) agarose gel and transferred to membranes (Hybond N^+, Amersham) by downward alkaline blotting (Koetsier et al., 1993). The membrane in Figure 2 was hybridized to a fluorescein-labeled DNA probe covering the complete SrMV-SCH CP-coding region. Labeling, hybridization, washing, and detection were performed according to the manufacturer's instructions (Gene Images, Amersham). The membrane in Figure 6 was hybridized to a ^{32}P-labeled DNA probe covering part of the SrMV-SCH CP-coding region. Labeling was done using a nick translation kit (GIBCO-BRL); hybridization and washing were performed according to previously published procedures (Sambrook et al., 1989).

RNA Gel-Blot Analysis

Total RNA was extracted from sugarcane leaves as previously described (Jones et al., 1985). RNA samples (15 μg of total RNA or the indicated amounts of poly(A^+) RNA) were electrophoresed on a 1.5% (w/v) agarose, 0.45 M formaldehyde, and 20 mM Hepes gel, pH 7.4, transferred to Hybond N^+ membranes by downward alkaline blotting (Ingelbrecht et al., 1998), and hybridized with a fluorescein-labeled DNA probe covering the complete SrMV-SCH CP-coding region. Labeling, hybridization, washing, and detection were performed according to the manufacturer's instructions (Gene Images, Amersham). Poly(A^+) RNA was isolated using the Oligotex mRNA Midi kit (Qiagen, Chatsworth, CA).

Isolation of Nuclei and Nuclear Run-Off Transcription Assays

Nuclei were isolated from fresh leaf tissue of field-grown sugarcane and the transcription assay and the conditions for prehybridization were as previously described (Ingelbrecht and de Carvalho, 1992). Denatured, double-stranded DNA fragments (approximately 0.25 μg per slot) or single-stranded phagemid-derived DNA (1 μg per slot) or single-stranded in vitro-generated RNA (0.5 μg per slot) were slot blotted onto a nitrocellulose membrane using a microfiltration apparatus (Bio-Rad), according to the manufacturer's instructions. Single-stranded DNA was prepared from phagemids using the helper phage M13KO7, according to standard procedures (Sambrook et al., 1989). In vitro transcription reactions were performed using an in vitro transcription system (Promega) with the following linearized plasmids: pIVING366-1 and -2 digested with *Not*I for antisense and sense *scrbcs-1* sequences respectively; pIVING362-3 and -4 digested with *Xho*I for antisense and sense CP RNA, respectively; the plasmid SCUBI561 containing a sugarcane polyubiquitin cDNA clone (Albert et al., 1995) was linearized with *Bam*HI and *Hin*dIII to produce antisense and sense ubiquitin RNA, respectively.

RESULTS

Transgene Constructs and Plant Transformation

The establishment and maintenance of a collection of seven SCMV strains that currently cause disease in sugarcane throughout the world was previously described (Yang and Mirkov, 1997). This collection includes the SCMV strains A, B, D, and E and the SrMV strains SCH, SCI, and SCM. For each strain, 2 kb of the 3′ terminus, encompassing the complete CP ORF, was sequenced. The nucleotide sequence identity of the CP genes within the strains of each group is more than 95%, whereas the nucleotide sequence identity between the two groups averages approximately 75%.

To engineer resistance in sugarcane, the CP-coding region and 77 nucleotides of the 234 3′UTR nucleotides of SrMV-SCH were fused to the maize ubiquitin-1 promoter with the first exon and intron and the 3′-terminator sequence of the *nopaline synthase* gene, yielding the plasmid *Ubi-hut*, illustrated in Figure 1. The *Ubi*-1 promoter region was previously shown to drive high levels of gene expression in sugarcane (Gallo-Meagher and Irvine, 1993). In

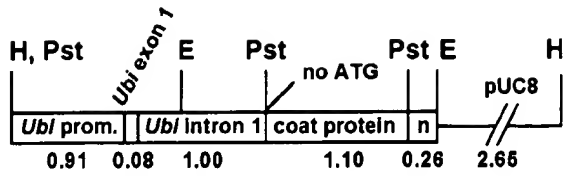


Figure 1. Schematic representation of the *Ubi-hut* plasmid (6.0 kb) linearized at the unique *Hin*dIII site showing the maize ubiquitin promoter (*Ubi* prom.), exon 1 and intron 1 plus the SrMV-SCH-untranslatable CP-coding region with 77 nucleotides of the 234 3'UTR sequence, and *nos* termination sequence (n) cloned into pUC8. Restriction sites: E, *Eco*RI; H, *Hin*dIII; Pst, *Pst*I. Numerical values indicate approximate kilobase pairs.

potyvirus-infected plants the CP gene product is produced by proteolytic cleavage of a polyprotein, and the CP ORF does not have a translation initiation codon at its 5'-proximal end. In the CP ORF the first start codon is located 17 nucleotides downstream of the CP 5'-proximal end, in a different reading frame from the CP ORF. The first start codon in-frame with the CP ORF is located approximately 313 nucleotides further downstream. Upon translation, this would yield a CP gene product that is truncated by about one-third (111 amino acids) of its normal size. A truncated CP gene product was never detected by ELISA or in western blots (data not shown).

The efficiency of cobombardment is very high in sugarcane; therefore, we used this strategy for generating transgenic plants expressing the *Ubi-hut* CP gene. Embryogenic callus derived from the commercial sugarcane varieties CP65-357 and CP72-1210 was bombarded with *Ubi-hut* in combination with *Ubi-bar* or *Ubi-npt* and selected on bialaphos- or geneticin-containing medium, respectively (see "Materials and Methods"). Approximately 220 plants were regenerated and transferred to the greenhouse for virus inoculations.

Transgene Integration Patterns

To assess the number of unique transformation events among the 220 plants, DNA gel-blot analysis was performed. Plant DNA was digested with *Hin*dIII and hybridized with the SrMV-SCH CP-coding region. In total, 29 different CP hybridization profiles were identified and the transgenics were grouped accordingly in groups numbered 1 through 29. More than one-half of the 220 plants belonged to only three groups and most of these redundant plants were discarded for practical reasons. The CP hybridization pattern of transgenic plants representing 26 groups appears in Figure 2. *Hin*dIII-digested DNA from at least one plant of each group was then hybridized to *bar* and/or *npt* gene sequences. Because *Hin*dIII cuts only once within the plasmid sequences, the copy number of the CP transgenes and the selectable marker genes can be estimated by scoring the number of bands on the DNA gel blots. All bands were scored once, irrespective of size and intensity of hybridization. The estimated gene copy numbers for the 29 groups are summarized in Table I; the plants shown in Figure 2 are listed in this table according to their group number.

As shown in Table I, the CP transgene copy number varies between 0 and more than 15, with about one-half of the plants having 4 to 9 copies. All plants selected on geneticin had both *npt* and CP sequences as shown in Figure 2. Insertion of only the *bar* selectable marker gene occurred once; plant 314 had no CP sequences but contained a single *bar* insertion. This indicates a high degree of cotransformation, in agreement with previous results on biolistic cotransformation of sugarcane and other monocots such as rice and barley (Wan and Lemaux, 1994; Bower et al., 1996; Qu et al., 1996). Note that plants 5 and 6 have a similar, yet clearly distinct, CP hybridization profile. Comparable observations were made for plants in groups 24 and 25 and also in groups 26 and 27 (data not shown). Plants with highly related transgene integration patterns most likely originated from the same transformation event and might represent chimeras for the CP transgene and/or the selectable marker gene.

Virus-Inoculated Transgenic Plants Display an Immune, Recovery, or Susceptible Phenotype That Is Not Strictly Correlated with Genotype

Transgenic plants were repeatedly inoculated with SrMV-SCH at the four- to eight-leaf stage, as described in "Materials and Methods." The plants were grown in the greenhouse for 8 to 10 months and then transplanted to the field where they were exposed to natural infection. The incidence of infection of nontransgenic controls in the field was approximately 30%.

Symptom development was periodically monitored by visual inspection and, in selected cases, also by ELISA or RNA gel-blot analysis. Three different responses were ob-

Figure 2. DNA gel-blot analysis of CP transgene integration patterns and copy number in sugarcane transgenics. DNA gel blot of total DNA (10 μg) digested with *Hin*dIII and hybridized with a fluorescein-labeled 1.1-kb *Eco*RI fragment containing the SrMV-SCH CP and 3'UTR. Numbers above the lanes refer to individual sugarcane transgenics. I, R, and S denote immune, recovered, and susceptible plants, respectively; N refers to plants that were not inoculated and with an undetermined phenotype. Transgenics were selected on bialaphos (*bar-hut*) or on geneticin (*npt-hut*), as indicated. DNA size markers are shown on the left.

Table I. *Summary of transgene copy number in transgenic sugarcane, as determined by hybridization of HindIII-digested DNA to the CP, bar, and npt probes*

Group No.	Copy No.			Plant No.
	CP	bar	npt	
bar-hut				
1	13	4	–	313
2	7	4	–	125
3	5	5	–	120
4	3	6	–	16
5	2	1	–	20
6	5	2	–	322
7	10	5	–	62
8	6	2–5	–	NR[d]
9	2	5	–	40
10	≥15	5	–	18
11	≥15	13	–	453
12	8	4	–	494
13	3	1–3	–	392
14	1	1	–	319
15	0	1	–	314
npt-hut				
16	5	–	≥15	459
17	4	–	6	41
18[a]	12	–	13	5
19[a]	10	–	13	6
20	4	–	2–4	35
21	13	–	≥15	32
22	5	–	10	460
23	6	–	13	477
24[b]	13	–	≥15	467
25[b]	14	–	≥15	NR
26[c]	7	–	13	463
27[c]	5	–	8	NR
28	3	–	5	33
29	8	–	13	479

[a,b,c] Groups with related transgene integration patterns. [d] NR, Not represented in the DNA gel blot of Figure 2.

served: an immune, a susceptible, and a recovery phenotype. Immune plants never developed symptoms even though some were inoculated up to six times. Susceptible and recovery plants developed mosaic symptoms 2 to 5 weeks after inoculation on the newly developing leaves. In susceptible plants these symptoms persisted and were most obvious on the younger leaf tissue. The recovery phenomenon was manifested by a gradual reduction in mosaic symptoms in successive leaves until leaves emerged that were completely symptomless and virus free. Resistant tissue from a recovery plant was phenotypically similar to immune tissue. In mature plants these two phenotypes could no longer be distinguished because of senescence of the older, symptomatic leaves in the recovered plants. The characteristic mosaic pattern of an inoculated, susceptible leaf is compared to an immune, symptomless leaf in Figure 3.

The time required for full recovery was variable. For example, recovery occurred relatively fast for plants in group 4; i.e. virus-free leaves appeared approximately 2 months after initial symptom development. However, symptoms persisted for approximately 1 year in plants 18 (group 10) and 313 (group 1) before recovery. Table II summarizes the phenotypes of individual transgenics from 25 different groups. Note that plants with the same transgene integration pattern did not necessarily display the same phenotype, as illustrated in Table II. Groups 2, 4, and 7 contain immune and recovery plants, whereas groups 1, 10, 17, and 20 include recovery and susceptible individuals. Group 6 includes immune, recovery, and susceptible individuals. Resistant plants tend to have an intermediate CP gene copy number, ranging from 4 to 10, whereas the number of CP insertions in most susceptible plants is either lower or higher.

Virus Resistance Requires a High Degree of Sequence Similarity

To determine whether virus resistance was strain specific, SrMV-SCH-resistant shoots were reinoculated with SrMV-SCH, with the closely related strains SrMV-SCI and -SCM, or with the more distantly related strain SCMV-D. Initial experiments had indicated that the resistance phenotype of group 4 transgenics, derived from either a recovered or immune plant, was maintained in shoots obtained after vegetative propagation. Stalks from different transformants of this group were cut into two- to four-bud sections and shoots were produced from these cuttings. Twenty-five shoots were sap inoculated with each of these virus strains.

As summarized in Table III, most shoots challenged with the more distantly related strain SCMV-D (22 of 25) devel-

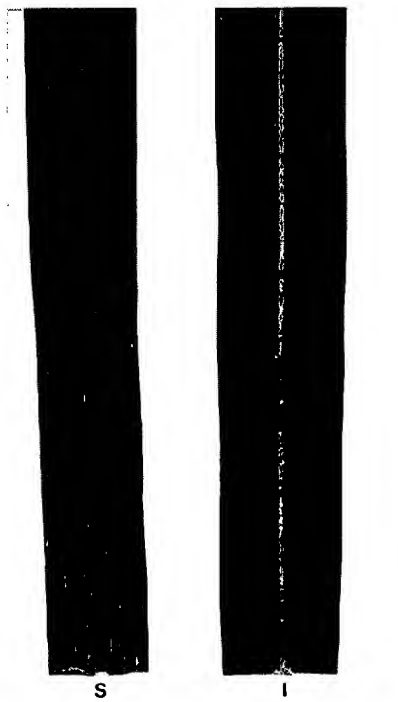


Figure 3. Phenotypes of SrMV-SCH-inoculated transgenic sugarcane plants. Leaves were taken from the same position on each of the plants. S, Susceptible leaf with typical mosaic symptoms; I, immune leaf without symptoms.

Table II. *Phenotypes of individual transgenic sugarcane plants*

Group No.	Plant No.			CP Copy No.
	Immune	Recovery	Susceptible	
3	120, 124			5
2	125	100, 103, 107, 119, 123		7
4	11, 16, 17	12, 13, 15, 23		3
7	60, 61, 62	3, 4, 38, 39, 57 58, 159		10
5		20		2
9		40		2
22		368		5
29		31		8
1		313	315	13
10		18	22	≥15
17		41, 46, 52, 56	48, 95	4
20		30	25, 35, 36	4
14			319	1
15			314	0
21			32	13
11			543	≥15
13			392	3
18			5, 24	12
19			6	10
23			359, 376, 469, 475, 476, 477	6
24			377	13
25			356	14
26			463	7
28			33	3
6	323	399, 424	322, 400, 525	5

oped symptoms. Symptom development was far less frequent after inoculation with the more closely related strains SrMV-SCI or -SCM (7 and 8 of 25, respectively). In SrMV-SCH challenge inoculations, 3 of 25 plants developed symptoms, indicating that the original resistance broke down at low frequency. However, these plants had completely recovered 3 months after inoculation, as demonstrated by RNA gel-blot analysis (data not shown). Symptomatic and recovered tissue was harvested from two of these plants and used for the study of DNA methylation patterns (see below).

Table III. *Response of SrMV-SCH-resistant shoots to challenge inoculations with SrMV and SCMV virus strains*

Inoculated Virus Strain	CP Sequence Identity[a]	Plants with Symptoms	
		Transgenic[b]	Nontransgenic control
	%	*no.*	
SrMV-SCH	100	3/25	8/10
SrMV-SCM	95	8/25	5/5
SrMV-SCI	95	7/25	3/3
SCMV-D	75	22/25	5/5

[a] Nucleotide sequence identity between the CP ORF of the transgene and the inoculated virus strain. [b] SrMV-SCH-resistant transformants from group 4 were vegetatively propagated by cuttings and shoots were sap inoculated with different SrMV and SCMV strains at the three- to five-leaf stage. Symptoms were visually scored 2 months after infection.

Most Resistant Plants Contain Reduced CP Transgene Steady-State mRNA Levels

To determine CP transgene and viral RNA levels in immune and susceptible plants, total RNA was extracted from leaves of selected 1-year-old, field-grown plants. When probed for CP sequences, the symptomatic plants 35, 463, and 525 showed a smear of RNA extending from approximately 10 kb, i.e. the size of the viral genomic RNA, to a few hundred bases, as shown in Figure 4A. This smear of viral RNA, which had been previously observed in potyvirus-infected material (Vance, 1991), prevented detection of the transgenic CP mRNA in the inoculated susceptible plants. As expected, no viral RNA was detected in the immune plants 17, 60, 120, and 323. The level of steady-state CP mRNA in plants 17, 60, and 120 is either below the detection limit or lower than in plant 323, which had a CP mRNA level comparable to that of the noninoculated susceptible plant 463 (Fig. 4B).

The RNA gel blots in Figure 4 show multiple CP RNA transcripts. The 1.3-kb mRNA, indicated with the larger arrowhead, corresponds to the predicted mRNA product of the *Ubi-hut* construct. The doublet indicated with the smaller arrowhead was observed in all plants with detectable levels of the 1.3-kb CP mRNA, including the noninoculated, susceptible plant 463 (Fig. 4B). These two bands were polyadenylated (Fig. 4B) and cross-hybridized with the ubiquitin intron 1 sequence (data not shown), suggesting that they represent incompletely processed transcripts. Some less-than-full-length, polyadenylated transcripts were specific for the immune plants 60 and 62, as shown in Figure 4B.

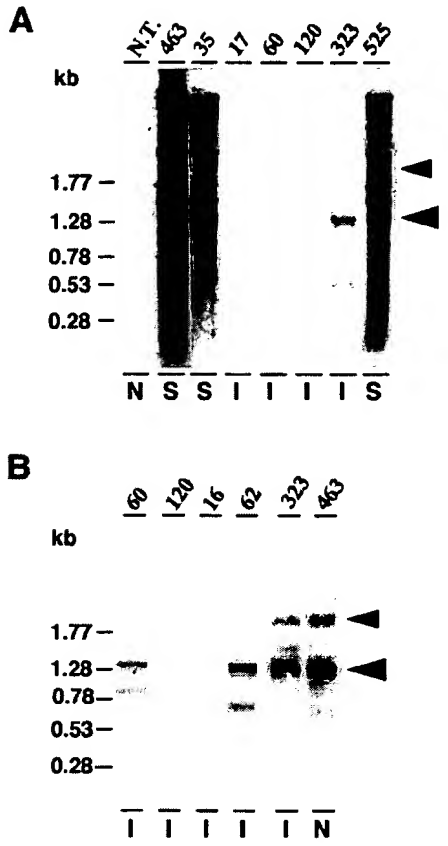


Figure 4. RNA gel-blot analyses of CP transgenic RNA and viral RNA levels in selected immune and susceptible transgenic sugarcane plants and a nontransgenic control. A, RNA gel blot of total RNA (15 μg per lane) hybridized with a fluorescein-labeled 1.1-kb *Pst*I fragment containing the SrMV-SCH CP ORF and 3'UTR. Numbers on top indicate sugarcane transgenics: N.T. is a nontransgenic, noninoculated (N) control; 463, 35, and 525 are susceptible (S) plants; 17, 60, 120, and 323 are immune (I) plants. RNA size markers are indicated on the left. The larger arrowhead on the right indicates the 1.3-kb CP mRNA; the smaller arrowhead indicates transcripts that cross-hybridize with the ubiquitin intron 1 sequence. B, RNA gel blot of poly(A$^+$) RNA from immune (I) plants 60, 120, 16, 62, and 323 and a noninoculated (N) susceptible control, 463; 0.9 μg of poly(A$^+$) RNA was loaded for plants 60 and 120; 2.5 μg for plants 16 and 62; and 0.8 μg for plants 323 and 463. RNA sizes are indicated on the left. Probes and hybridization were as in A.

CP Transgenes in Resistant and Susceptible Plants Are Actively Transcribed

To compare transcriptional activity of the transgenes in susceptible and resistant plants, nuclei were isolated from the same leaf material that was used for the RNA gel-blot analyses in Figure 4. In vitro run-off transcription reactions were performed, and labeled nRNA was hybridized to slot-blotted DNA containing CP, *bar*, and *npt* gene sequences. Transcription along the 2.65-kb vector sequences downstream of the chimeric genes was also measured. The endogenous polyubiquitin genes served as an internal control.

As illustrated in Figure 5A, the CP transgenes were transcribed in all transgenic plants but not in the nontransgenic control. The *bar* gene was transcribed in plants 17, 60, 120, and 525 and the *npt* gene was transcribed in plant 35, as expected. Overall, the transcription rates of the CP transgenes correlated well with their estimated copy number (Table I). The immune phenotype was not correlated with a higher CP transcription rate. Together, Figures 4 and 5A show that the CP transgenes in the immune plants 17, 60 and 120 were actively transcribed, whereas their steady-state CP mRNA level was low or below the detection limit. This suggests that posttranscriptional degradation of the CP transcripts occurs, in agreement with a gene-silencing-related, virus-resistance mechanism.

Initial run-off assays using phagemid-derived single-stranded DNA showed strong hybridization signals with phagemids that did not carry an insert. This "background" hybridization was consistently observed in transgenic plants but not in a nontransgenic control (data not shown). There is sequence similarity between pUC8 and the phagemid, which suggested that the pUC8 vector sequences were

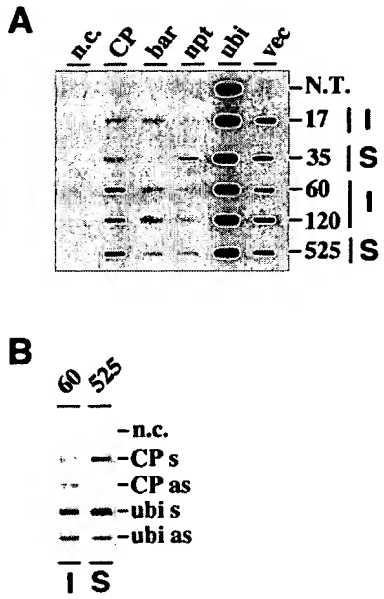


Figure 5. Analyses of transcription rates for selected immune and susceptible transgenic sugarcane plants and a nontransgenic control. Nuclei were isolated from leaves of 1-year-old, field-grown plants and used in nuclear run-off assays. A, Autoradiogram showing ^{32}P-labeled nRNAs hybridized to DNA fragments containing gene-specific sequences: CP, 1.1-kb coding sequence of the SrMV-SCH CP; bar, 0.58-kb coding sequence of the *bar* gene; npt, 0.76-kb 3' part of the *npt*-coding sequence; ubi, 0.74-kb 5' part of the SCUBI561 sugarcane polyubiquitin cDNA; vec, the 2.65-kb pUC8 sequences present downstream of the chimeric CP and marker genes; n.c., negative control (no DNA). Numbers on the right refer to plants used as a source of nuclei: 17, 60, and 120 are immune (I); 35 and 525 are susceptible (S) plants; N.T. is a nontransgenic, noninoculated control. B, Autoradiogram showing ^{32}P-labeled nRNAs of the immune plant 60 and the susceptible plant 525 hybridized to in vitro-synthesized single-stranded RNA. Abbreviations are as in A. Sense (s) and antisense (as) refers to the polarity of the nascent RNA.

actively transcribed. Figure 5A confirms that the 2.65-kb vector sequences present downstream of the transgenes were indeed transcribed at similar rates in all transgenic plants.

In Figure 5B, in vitro-transcribed, single-stranded RNA was bound to the membrane. As shown, both sense and antisense strands of the CP transgenes were transcribed to an equal extent in the immune plant 60. Antisense transcription of the CP transgenes also occurred in the susceptible plant 525 but at a much lower rate. Surprisingly, antisense transcription was also evident for the endogenous ubiquitin genes.

Specific Sites Located in the Transcribed Region of the CP Transgenes Are More Extensively Methylated in Most Resistant Plants

To examine whether virus resistance was correlated with increased methylation of the CP transgenes, we performed DNA gel-blot analyses using total DNA isolated from resistant and susceptible leaf tissue.

This analysis was first performed on isogenic leaf material with and without symptoms. We used two recovered plants obtained after vegetative propagation and reinoculation with SrMV-SCH for this purpose (see above): one was derived from transformant 13, which originally had a recovery phenotype, and one was from transformant 16, which originally was immune. These two plants were showing clear symptoms and accumulated viral CP, as determined by ELISAs 2 months after inoculation. Four weeks later, both plants had developed new, symptomless leaves that were free of viral RNA, as judged by RNA gel-blot analysis (data not shown). Total DNA was isolated from leaf tissue with and without symptoms from these two plants. As a control, DNA was also isolated from two clones that did not develop symptoms upon reinoculation; one clone was derived from plant 13 and one was from plant 16. This DNA was digested with *Eco*RI and *Pvu*II, both of which are sensitive to cytosine methylation, and probed with a CP gene-specific fragment, illustrated in Figure 6A. Figure 6B indicates that the CP genes are more extensively methylated in recovered (R and R_1) or immune (I) tissue than in leaves with symptoms (S), as judged by the presence of the higher-M_r bands in lanes R, I, and R_1 versus the S lanes.

We then compared the methylation status of the same sites among independently transformed resistant and susceptible plants. As shown in Figure 6B, most immune plants displayed the 1.55- and/or the 1.95-kb bands in addition to the 1.3-kb band, indicating partial methylation of one or both the *Pvu*II sites. By contrast, only the 1.3-kb band was observed in the sensitive plants and in one immune plant, suggesting that these sites are not detectably methylated in these plants. The origin of the intense 1.8-kb band in lanes 13, 16, 17, and 120 is unclear. It might represent a junction fragment between plant and plasmid DNA or rearranged CP gene sequences. Using the same plant material, we could not detect methylation of the 10 *Sau*3AI and 4 *Hpa*II sites that were located in the 5′-ubiquitin promoter sequences (data not shown).

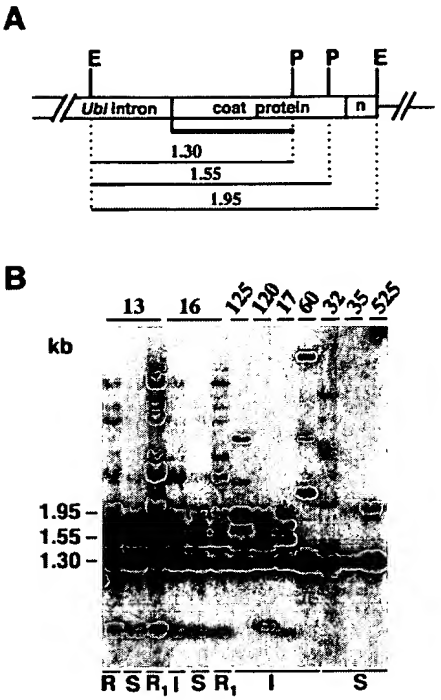


Figure 6. DNA gel-blot analysis of differential methylation patterns in total DNA from resistant and susceptible leaf tissue. A, Diagram showing the position of the *Eco*RI (E) and *Pvu*II (P) sites within the CP transgene. The solid line represents the 0.75-kb probe used for hybridization of the blot shown in B. The 1.30-kb fragment indicates complete digestion (no methylation) of the *Pvu*II and *Eco*RI sites. The 1.55- and 1.95-kb fragments indicate partial digestion (partial methylation) of the *Pvu*II site(s). B, DNA gel blot of total DNA (10 μg) digested with *Eco*RI and *Pvu*II and hybridized with the probe shown in A. Numbers refer to sugarcane transgenics: 125, 120, 17, and 60 are immune (lane I); 32, 35 and 525 are susceptible (lane S). For plants 13 and 16: lanes S and R_1 indicate symptomatic and resistant tissue harvested from the same plant before and after recovery; lanes R and I indicate resistant tissue from a different 13 and 16 clone that did not show symptoms upon re-inoculation. Hybridizing bands that correspond to the fragments shown in A are indicated on the left.

DISCUSSION

Resistance to SrMV-SCH was obtained in transgenic sugarcane plants expressing an untranslatable form of the SrMV-SCH CP gene. Several lines of evidence indicate that the underlying resistance mechanism in the investigated clones is related to PTGS.

In general, transgenic sugarcane plants challenged with SrMV-SCH fell into three phenotypic classes: (a) plants that were susceptible and showed symptoms throughout the 2-year monitoring period; (b) recovery plants, i.e. plants that initially showed symptoms but, during further growth, developed symptomless leaves that were virus free; and (c) immune plants that did not show symptoms and were virus free (Table II). The time required for recovery was variable; it was relatively fast in group-4 plants, which were completely virus free 2 to 3 months after initial symptom development, but not in plants 18 and 313, in which symptoms persisted for approximately 1 year.

SrMV-SCH-resistant shoots were resistant to challenge inoculations with SrMV-SCH and the closely related SrMV-SCI and -SCM strains (95% nucleotide similarity) but not to inoculation with the more distantly related SCMV strain D (75% nucleotide similarity, Table III). The immune and recovery phenotype and the strain-specific resistance are well-documented characteristics of gene-silencing-related, virus-resistance phenomena in dicots (Lindbo et al., 1993; Smith et al., 1994; Mueller et al., 1995; Tenllado et al., 1995, 1996; Goodwin et al., 1996; Pang et al., 1996; Prins et al., 1996). The inoculation experiments also demonstrated that shoots obtained after vegetative propagation of immune or recovered plants remained resistant. The fact that sugarcane is a vegetatively propagated crop may therefore prove advantageous in maintaining valuable clones, because PTGS can be reversed during meiosis.

In total, 220 transgenic plants were regenerated; they could be classified into 29 different groups based on their CP and marker-transgene-integration patterns (Fig. 1; Table I). Transformants with the same CP transgene integration pattern did not necessarily display the same phenotype upon inoculation. In fact, five groups had both recovered and susceptible individuals (Table II), and one of these groups (group 6) even included an immune individual. It has been previously reported that isogenic transgenic lines can display different levels of resistance upon virus inoculation in dicots (Smith et al., 1994; Sijen et al., 1996). This differential response occurred more frequently when the length of the homologous sequence between the transgene and the inoculated virus was reduced (Sijen et al., 1996). This result is comparable to what we observed in our inoculation studies using different virus strains. When SrMV-SCH resistant plants were infected with a virus strain with decreasing levels of sequence homology, from 100% to 95% to 75%, an increasing number of plants developed symptoms, from 3 to 7 or 8 to 22 (of 25), indicating a positive correlation between the level of resistance and the sequence homology between the CP ORF of the infecting virus and the transgene. This result further suggests that the transition from a susceptible to a resistant phenotype is gradual and continuous. As a whole, we can conclude that the resistance mechanism in sugarcane holds quantitative aspects and can operate with variable efficiency in plants with certain genetic backgrounds.

On RNA gel blots, the susceptible plants 35 and 525 showed a smear of viral RNA that obscured the transgene-derived CP mRNA (Fig. 4). No viral RNA could be detected in the immune plants 17, 60, and 120, and the steady-state transgene CP mRNA level in these plants was undetectable or reduced, compared with the noninoculated susceptible control 463, which is in agreement with a PTGS-related, virus-resistance mechanism (Lindbo et al., 1993; Smith et al., 1994; Mueller et al., 1995; Pang et al., 1996; Sijen et al., 1996). The steady-state CP mRNA level in the immune plant 323 was higher than in the other immune plants and similar to that in the noninoculated susceptible plant 463. Perhaps the resistance mechanism in this plant is not related to PTGS. On the other hand, Tenllado et al. (1995) and Marano and Baulcombe (1998) described putative examples of HDR where resistant plants retained high levels of the homologous transgenic mRNA. RNA gel-blot analyses revealed a number of transcripts in addition to the expected 1.3-kb CP mRNA. Two polyadenylated transcripts of approximately 2.4 kb were observed in all plants with detectable levels of the 1.3-kb CP mRNA. These bands cross-hybridize with the *Ubi*-1 intron sequence, suggesting that they represent incompletely processed CP transcripts. Because these bands also occur in the noninoculated susceptible control 463, they are not related to transgene silencing. Some less than full-length, polyadenylated bands were detected in the immune plants 60 and 62 but not in the noninoculated, susceptible control 463 (Fig. 4B). Truncated transcripts have previously been shown in other cases of HDR (Goodwin et al., 1996; Tanzer et al., 1997) and may represent intermediates of the RNA-degradation mechanism.

The immune plants 17, 60, and 120 having low or undetectable CP steady-state mRNA levels display active transcription of the CP transgenes in the nucleus, as shown in nuclear run-off assays (Fig. 5). This result suggests that the CP mRNA level in these plants is down-regulated by a posttranscriptional mechanism, as expected for HDR (Lindbo et al., 1993; Mueller et al., 1995; Prins et al., 1996; Sijen et al., 1996). There was no obvious correlation between transcription rates of the CP transgenes and an immune or susceptible phenotype. In the immune plant 60, both sense and antisense transcription of the CP transgenes occurred at similar rates. Antisense transcription was also detected for the CP genes of the susceptible plant 525 but at a much lower rate than sense transcription. Most surprisingly, we also detected antisense transcription for the endogenous ubiquitin genes. Although highly unexpected, this finding might not be unprecedented, because a previous report (Christensen and Quail, 1989) described antisense transcription for the maize ubiquitin genes and possibly also for maize rDNA. Whether antisense transcription of these host genes should be considered background, as suggested by these authors, or indicates a true event is not clear at present. Finally, the run-off assays clearly showed that the pUC8 vector sequences located downstream of the CP and/or marker genes were actively transcribed in all transgenic plants. Fortuitous transcription of vector sequences might result from transcriptional readthrough originating from the transgenic *Ubi*-1 promoter or from endogenous plant promoters that flank the integrated transgenes. The presence of truncated or rearranged plasmid sequences, often observed in biolistic transformation, possibly enhanced this unexpected transcription.

A correlation between methylation of transgenes in the transcribed region and resistance has been found in several examples of HDR in dicots (Smith et al., 1994; English et al., 1996; Sijen et al., 1996) as well as in our study. Using leaf material with and without symptoms isolated from the same plant before and after recovery, we could demonstrate that specific sites located in the coding region of the CP transgenes were more extensively methylated in resistant tissue (Fig. 6). Methylation of the CP transgenes was most pronounced in tissue that had just recovered, suggesting that DNA methylation occurs primarily during the initial establishment of resistance. When comparing inde-

pendently transformed susceptible and resistant plants, increased methylation of these sites was found in most but not all resistant plants. We could not detect clear methylation of sites located in the *Ubi*-1 promoter in resistant or susceptible tissue.

Various models have been proposed to explain PTGS and how it is involved in virus resistance, including the RNA threshold model (for review, see Dougherty and Parks, 1995) and the ectopic pairing model (for review, see Baulcombe and English, 1996). In essence, these models propose that enhanced turnover of transgene and viral RNA is mediated by cRNA molecules that are synthesized by a plant-encoded, RNA-dependent RNA polymerase. The trigger for turnover is unknown, but Dougherty and Parks (1995) proposed that this cRNA is synthesized in response to the presence of overexpressed RNA molecules. According to the ectopic pairing model, specific DNA-DNA interactions result in the formation of aberrant RNA, which is copied into cRNA and leads to the degradation of homologous RNA molecules. Recently, Waterhouse et al. (1998) proposed an interesting variant of these models. These authors suggested that PTGS is induced by double-stranded RNA, which serves as a template for the production of cRNA by the RNA-dependent RNA polymerase. Our run-off data do not support a model that is based strictly on exceeding a critical level of RNA because the transcription rate of the CP transgenes in immune plants is not always higher than in susceptible transgenics. As mentioned, in the immune plant 60, the CP transgenes were transcribed in both directions at similar rates. This could potentially lead to the formation of double-stranded RNA, consistent with the model of Waterhouse et al. (1998). However, we could not detect CP antisense transcripts in total RNA preparations from this plant, as judged by RNA gel-blot analysis (data not shown).

In a recent review of epigenetic transgene silencing, Matzke and Matzke (1998) pointed out that the genomes of agriculturally important cereal crops are often substantially more complex than that of Arabidopsis or of the *Nicotiana* species, which are generally used in these studies. These authors then raised the issue whether or not the variations in base composition between dicots and monocots could be responsible for different types of gene silencing. Sugarcane has a particularly complex genome, because it is an allopolyploid with variable chromosome numbers and ploidy levels. Despite these differences, we have also recognized the hallmarks of PTGS and HDR in dicot plants in our study. Our data provide evidence that an ancestral pathway for posttranscriptional gene regulation and RNA surveillance is conserved between monocots and dicots. This is consistent with reports of related gene-silencing mechanisms in fungi (for review, see Cogoni and Macino, 1997), *Paramecium* (Ruiz et al., 1998), and *Caenorhabditis elegans* (Fire et al., 1998). Still, we are only beginning to understand these phenomena; studies at the molecular level may reveal differences between PTGS and HDR in monocots versus dicots. The recent identification of a viral suppressor of PTGS in tobacco etch virus offers approaches to directly address these questions. From a more applied perspective, our data demonstrate the feasibility of this approach to engineering virus resistance in this agronomically important group of plants.

ACKNOWLEDGMENTS

We would like to thank Josefina Bustamante, Mercedes Campos, Teresa De La Garza, and Eduardo Hernandez for plant cell tissue culture and plant maintenance. We are appreciative to Dr. F. Moonan for helpful discussions throughout this work and to Dr. Z.N. Yang for help with constructing the plasmids *Ubi-npt* and *Ubi-hut*. We are grateful to Dr. H. Albert and Dr. W. Tang for providing the plasmids SCUBI561 and *scrbcs*-1.

Received October 30, 1998; accepted January 7, 1999.

LITERATURE CITED

Albert HH, Carr JB, Moore PM (1995) Nucleotide sequence of a sugarcane polyubiquitin cDNA (accession no. L41658) (PGR 95–045). Plant Physiol **109:** 337

Al-Kaff NS, Covey SN, Kreike MM, Page AM, Pinder R, Dale PJ (1998) Transcriptional and posttranscriptional plant gene silencing in response to a pathogen. Science **279:** 2113–2115

Anandalakshmi R, Pruss GJ, Ge X, Marathe R, Mallory AC, Smith TH, Vance VB (1998) A viral suppressor of gene silencing in plants. Proc Natl Acad Sci USA **95:** 13079–13084

Arencibia A, Molina P, de la Riva G, Selman-Housein G (1995) Production of transgenic sugarcane (*Saccharum officinarum* L.) plants by intact cell electroporation. Plant Cell Rep **14:** 305–309

Baulcombe DC (1996) Mechanisms of pathogen-derived resistance to viruses in transgenic plants. Plant Cell **8:** 1833–1844

Baulcombe DC, English JJ (1996) Ectopic pairing of homologous DNA and post-transcriptional gene silencing in transgenic plants. Curr Opin Biotechnol **7:** 173–180

Bower R, Birch RG (1992) Transgenic sugarcane plants via microprojectile bombardment. Plant J **2:** 409–416

Bower R, Elliott AR, Potier BAM, Birch RG (1996) High-efficiency, microprojectile-mediated co-transformation of sugarcane, using visible or selectable markers. Mol Breed **2:** 239–249

Burner DM, Legendre BL (1994) Cytogenetic and fertility characteristics of elite sugarcane clones. Sugar Cane **1:** 6–10

Christensen AH, Quail PH (1989) Sequence analysis and transcriptional regulation by heat shock of polyubiquitin transcripts from maize. Plant Mol Biol **12:** 619–632

Christensen AH, Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res **5:** 213–218

Cogoni C, Macino G (1997) Conservation of transgene-induced post-transcriptional gene silencing in plants and fungi. Trends Plant Sci **11:** 438–443

Converse RH, Martin RR (1990) ELISA methods for plant viruses. *In* R Hamptom, E Ball, S De Boer, eds, Serological Methods for Detection and Identification of Viral and Bacterial Plant Pathogens. APS Press, St. Paul, MN, pp 179–196

Covey SN, Al-Kaff NS, Lángara A, Turner DS (1997) Plants combat infection by gene silencing. Nature **385:** 781–782

Depicker A, Van Montagu M (1997) Post-transcriptional gene silencing in plants. Curr Opin Cell Biol **9:** 373–382

Dougherty WG, Parks TD (1995) Transgenes and gene suppression: telling us something new? Curr Opin Cell Biol **7:** 399–405

English JJ, Davenport GF, Elmayan TE, Vaucheret H, Baulcombe DC (1997) Requirement of sense transcription for homology-dependent virus resistance and trans-inactivation. Plant J **12:** 597–603

English JJ, Mueller E, Baulcombe DC (1996) Suppression of virus accumulation in transgenic plants exhibiting silencing of nuclear genes. Plant Cell **8:** 179–188

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature **391:** 806–811

Gallo-Meagher M, Irvine JE (1993) Effects of tissue type and promoter strength on transient GUS expression in sugarcane following particle bombardment. Plant Cell Rep **12:** 666–670

Gallo-Meagher M, Irvine JE (1996) Herbicide resistant transgenic sugarcane plants containing the *bar* gene. Crop Sci **36:** 1367–1374

Goodwin J, Chapman K, Swaney S, Parks TD, Wernsman EA, Dougherty WG (1996) Genetic and biochemical dissection of transgenic RNA-mediated virus resistance. Plant Cell **8:** 95–105

Ingelbrecht I, de Carvalho F (1992) Isolation of nuclei and *in vitro* run-off transcription. *In* D Inzé, D Van Der Straeten, M Van Montagu, eds, Practical Course on Plant Molecular Biology. EMBO, Laboratorium voor Genetica, Gent, Belgium, pp 117–132

Ingelbrecht IL, Mandelbaum CI, Mirkov TE (1998) Highly sensitive northern hybridization using a rapid protocol for downward alkaline blotting of RNA. BioTechniques **25:** 420–425

Irvine JE (1999) *Saccharum* species as horticultural classes. Theor Appl Genet (in press)

Jones JDG, Dunsmuir P, Bedbrook J (1985) High level expression of introduced chimeric genes in regenerated transformed plants. EMBO J **4:** 2411–2418

Kasschau KD, Carrington JC (1998) A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. Cell **95:** 461–470

Koetsier PA, Schorr J, Doerfler W (1993) A rapid optimized protocol for downward alkaline southern blotting of DNA. BioTechniques **15:** 260–262

Lindbo JA, Silva-Rosales L, Proebsting WM, Dougherty WG (1993) Induction of a highly specific antiviral state in transgenic plants: implications for regulation of gene expression and virus resistance. Plant Cell **5:** 1749–1759

Marano MR, Baulcombe DC (1998) Pathogen-derived resistance targeted against the negative-strand RNA of tobacco mosaic virus: RNA strand-specific gene silencing? Plant J **13:** 537–546

Matzke MA, Matzke AJM (1998) Epigenetic silencing of plant transgenes as a consequence of diverse cellular defense responses. CMLS Cell Mol Life Sci **54:** 94–103

Mueller E, Gilbert J, Davenport G, Brigneti G, Baulcombe DC (1995) Homology-dependent resistance: transgenic virus resistance in plants related to homology-dependent gene silencing. Plant J **7:** 1001–1013

Palauqui J-C, Elmayan T, Pollien J-M, Vaucheret H (1997) Systemic acquired silencing: transgene-specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions. EMBO J **16:** 4738–4745

Pang SZ, Jan F-J, Carney K, Stout J, Tricoli DM, Quemada HD, Gonsalves D (1996) Post-transcriptional transgene silencing and consequent tospovirus resistance in transgenic lettuce are affected by transgene dosage and plant development. Plant J **9:** 899–909

Prins M, Resende RD, Anker C, van Schepen A, de Haan P, Goldbach R (1996) Engineered RNA-mediated resistance to tomato spotted wilt virus is sequence specific. Mol Plant–Microbe Interact **9:** 416–418

Qu R, De Kochko A, Zhang L, Marmey P, Li L, Tian W, Zhang S, Fauquet CM, Beachy RN (1996) Analysis of a large number of independent transgenic rice plants produced by the biolistic method. In Vitro Cell Dev Biol Plant **32:** 233–240

Ratcliff F, Harrison BD, Baulcombe DC (1997) A similarity between viral defense and gene silencing in plants. Science **276:** 1558–1560

Ruiz F, Vayssié L, Klotz C, Sperling L, Madeddu L (1998) Homology-dependent gene silencing in *Paramecium*. Mol Biol Cell **9:** 931–943

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Shukla DD, Ward CW, Brunt AA (1994) The sugarcane mosaic virus subgroup. *In* The Potyviridae. CAB International, Wallingford, UK, pp 360–371

Sijen T, Wellink J, Hiriart J-B, van Kammen A (1996) RNA-mediated virus resistance: role of repeated transgenes and delineation of targeted regions. Plant Cell **8:** 2277–2294

Smith HA, Swaney SL, Parks TD, Wernsman EA, Dougherty WG (1994) Transgenic plant virus resistance mediated by untranslatable sense RNAs: expression, regulation, and fate of nonessential RNAs. Plant Cell **6:** 1441–1453

Stam M, Mol JNM, Kooter JM (1997) The silence of genes in transgenic plants. Ann Bot **79:** 3–12

Sugar and Sweetener–Situation and Outlook Yearbook (1997) U.S. Department of Agriculture, Washington, DC

Tai TH, Tanksley SD (1990) A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. Plant Mol Biol Rep **8:** 297–303

Tang W, Sun SSM (1993) Sequence of a sugarcane ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit gene. Plant Mol Biol **21:** 949–951

Tanzer MM, Thompson WF, Law MD, Wernsman EA, Uknes S (1997) Characterization of post-transcriptionally suppressed transgene expression that confers resistance to tobacco etch virus infection in tobacco. Plant Cell **9:** 1411–1423

Tenllado F, García-Luque I, Serra MT, Díaz-Ruíz JR (1995) *Nicotiana benthamiana* plants transformed with the 54-kDa region of the pepper mild mottle tobamovirus replicase gene exhibit two types of resistance responses against viral infection. Virology **211:** 170–183

Tenllado F, García-Luque I, Serra MT, Díaz-Ruíz JR (1996) Resistance to pepper mild mottle tobamovirus conferred by the 54-kDa gene sequence in transgenic plants does not require expression of the wild-type 54-kDa protein. Virology **219:** 330–335

Vance VB (1991) Replication of potato virus X is altered in coinfections with potato virus Y. Virology **182:** 486–494

Voinnet O, Baulcombe DC (1997) Systemic signalling in gene silencing. Nature **389:** 553

Wan Y, Lemaux PG (1994) Generation of large numbers of independently transformed fertile barley plants. Plant Physiol **104:** 37–48

Waterhouse PM, Graham MW, Wang M-B (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc Natl Acad Sci USA **95:** 13959–13964

Yang ZN, Mirkov TE (1997) Sequence and relationships of sugarcane mosaic and sorghum mosaic virus strains and development of RT-PCR-based RFLPs for strain discrimination. Phytopathology **87:** 932–939